NON-UNIFORMLY HYPERBOLIC HORSESHOES ARISING FROM BIFURCATIONS OF POINCARÉ HETEROCLINIC CYCLES

by JACOB PALIS *and* JEAN-CHRISTOPHE YOCCOZ*

ABSTRACT

In the present paper, we advance considerably the current knowledge on the topic of bifurcations of heteroclinic cycles for smooth, meaning C^∞, parametrized families $\{g_t \mid t \in \mathbf{R}\}$ of surface diffeomorphisms. We assume that a quadratic tangency q is formed at $t = 0$ between the stable and unstable lines of two periodic points, not belonging to the same orbit, of a (uniformly hyperbolic) horseshoe K (see an example at the Introduction) and that such lines cross each other with positive relative speed as the parameter evolves, starting at $t = 0$ and the point q. We also assume that, in some neighborhood W of K and of the orbit of tangency $o(q)$, the maximal invariant set for $g_0 = g_{t=0}$ is $K \cup o(q)$, where $o(q)$ denotes the orbit of q for g_0. We then prove that, when the Hausdorff dimension HD(K) is bigger than one, but not much bigger (see (H.4) in Section 1.2 for a precise statement), then for most t, $|t|$ small, g_t is a non-uniformly hyperbolic horseshoe in W, and so g_t has no attractors in W. Most t, and thus most g_t, here means that t is taken in a set of parameter values with Lebesgue density one at $t = 0$.

CONTENTS

1. Introduction

1.1. *The context.* — One of the most challenging problems in the theory of dynamical systems is to understand some of the main features, like creation of attractors, of the orbit structure of the dynamics arising from bifurcations of homoclinic or heteroclinic cycles. Typically, the cycles we consider display an orbit of non-transversal intersection between some stable and unstable manifolds of fixed or periodic orbits, whose unfolding leads to dynamics with a rich orbit structure. For surface diffeomorphisms, such orbit of non-transversal intersection of stable and unstable lines correspond to a homoclinic or

* Partially support by CNPq and FAPERJ, Brazil.

DOI 10.1007/s10240-009-0023-x

heteroclinic tangency. The existence of cycles was much emphasized by Poincaré in late 19th century, in his classic Nouvelles Méthodes de la Mécanique Céleste [Po], where he stated that “rien n’est plus propre à nous donner une idée de la complication de tous les problèmes de dynamique”. The question of bifurcations of cycles often arises when we consider parametrized families of dynamics, like in the present work: they are created and then bifurcate, as the parameter evolves.

In the case of surface diffeomorphisms, the richness of the dynamics obtained from unfolding a homoclinic or heteroclinic tangency is implicit in the work of Cartright-Littlewood [CL] more than sixty years ago and thereafter in several articles by the authors, especially the latter; see Levi [L] for an account of some explicit consequences of these works. In between, in the early sixties, and along the same line, a new and fundamental dynamical structure was exhibited by Smale [S]: the horseshoe map associated to a transversal homoclinic orbit. Besides that, Levi also made use of the following remarkable result of Newhouse [N]: under some mild conditions, the unfolding of a homoclinic tangency for C^r, $r \geq 2$, surface diffeomorphisms leads, in the C^r topology, to open sets of diffeomorphisms such that none of its elements is (uniformly) hyperbolic. That is, hyperbolic diffeomorphisms are not dense in the set of all such maps, which had been an important conjecture by Smale. He also showed that such an unfolding leads to the existence of open sets of diffeomorphisms with a dense (actually, Baire second category) subset of elements displaying each of them infinitely many simultaneous periodic attractors (sinks) or repellors (sources). We refer especially to [BDV], for a comprehensive presentation of the concepts and results that we have just mentioned.

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 pioneering work of Benedicks-Carleson [BC], and those of Mora-Viana [MV] and Colli [C]. 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] and [P3]) concerning the orbit structure of 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. Several other conjectures were formulated in the above works that together compose a global scenario for dynamics.

As mentioned in the Abstract, the present paper represents a contribution to the understanding of the dynamics arising from bifurcating a cycle of a C^∞ surface diffeomorphism. We consider one-parameter families of diffeomorphisms g_t containing the initial bifurcating diffeomorphism, 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 neighbor-

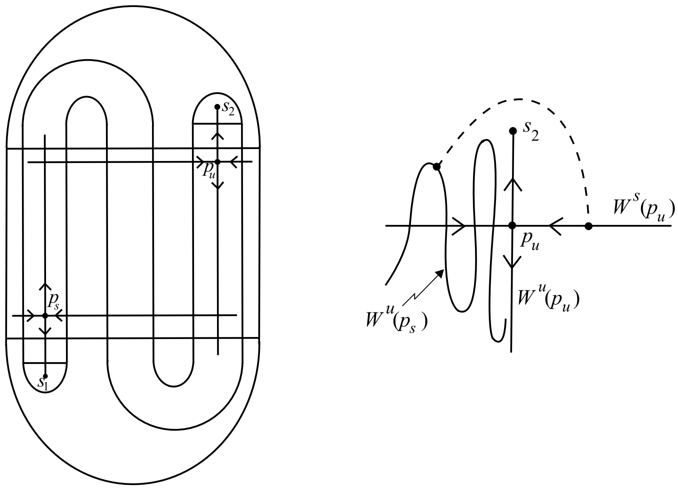

FIG. 1. — Horseshoe and homoclinic tangency

hood W 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 one. We show that right after the bifurcation, i.e. for $t > 0$ small, most diffeomorphisms are non-uniformly hyperbolic in W and so they display no attractors nor repellors in the neighborhood W of $K \cup o(q)$. This means that the parameter values corresponding to diffeomorphisms displaying no attractors nor repellors should have total density, that is density one, at $t = 0$. The concept is again discussed in the next subsection.

An example of the creation of a heteroclinic cycle associated to a (hyperbolic) horseshoe is indicated in Figure 1. Initially we have the 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 repellor outside the figure and two fixed point 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 as 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.

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 results to be true for all cases $0 < \mathrm{HD}(K) < 2$. To achieve 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 expect our results to be true in higher dimensions (see [MPV]). Finally, we hope that the ideas introduced in the present paper might be useful in broader contexts. In the horizon lies a famous question concerning the standard family of area preserving maps (see [BDV]): we ask whether we 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.

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 repellor. A horseshoe is topologically a Cantor set.

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

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

Observe that the smaller set

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

is a horseshoe which is the hyperbolic continuation of K.

Let $\mathcal{U}$ be an appropriately small neighbourhood of f in $\mathrm{Diff}^\infty(M)$. 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

$$(\mathbf{1.3}) \qquad \Lambda_g = K_g.$$

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

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

where q_g is the tangency point close to q given by the definition of $\mathcal{U}_0$. In the subsequent sections, we will omit the dependence on g in the notation and just write q for q_g when $g \in \mathcal{U}_0$.

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

$$\lim_{t\to 0} \frac{1}{t} \operatorname{Leb}(s \in (0,t], g_s \in \mathcal{P}) = 1. \tag{1.5}$$

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

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 repellors. It will be (non-uniformly) hyperbolic

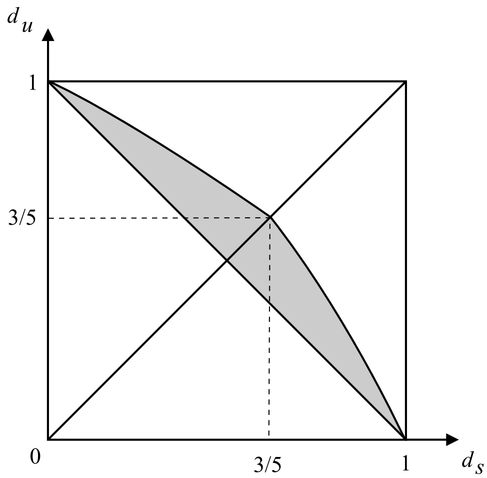

FIG. 2. — The dimension condition

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 Sections 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.

Remark **1.** — 1. 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.

2. The properties that we are proving for non-uniformly hyperbolic horseshoes are also true for uniformly hyperbolic ones, and indeed it is possible for Λ_g to be uniformly hyperbolic for a positive proportion of parameters even when $d_s^0 + d_u^0 \geqslant 1$. Therefore we will assume in the following that $d_s^0 + d_u^0 \geqslant 1$.

3. The tools that we develop probably allow to give, after some more work, further geometric and dynamical information on Λ_g (for most $g \in \mathcal{U}_+$) beyond that given in Sections 10 and 11. This could be the subject of subsequent investigations.

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 $R := \bigsqcup R_a$ 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

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

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 the 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 Section 5.4. The intuitive notion behind the formal definition is the following: an element (P, Q, n) (with Q crossing the domain L_u of G) should be I-transverse to an element (P', Q', n') (with P' crossing the image L_s of G) if the distance $\delta(Q, P')$ between the tip of the parabolic-like strip $G(Q \cap L_u)$ and P' satisfies

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

for all $t \in I$, where $|Q|$, $|P'|$ are the widths of the strips Q and P'. Actually, as the distance $\delta(Q, P')$ is expected to vary with the parameter with derivative close to one, and we want a uniform control in I, it is more natural to ask that

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

for all $t \in I$. 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 formal definition of the transversality relation that is more complicated than this. In Appendix C, we explain why this seemed complication is rather necessary. We use the notation $Q \pitchfork_I P'$ to say that Q is I-transverse to P'.

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. Thus, the construction of $\mathcal{R}(I_0)$ is easy. For smaller parameter intervals, what is clear from the defining axioms (R1)–(R7) is that there can be at most one class $\mathcal{R}(I)$ satisfying them. That such a class exists (for candidate intervals whose parent is regular) is stated in Section 5.4 but will only be completely proven at the end of Section 7.

We conclude Section 5 with the introduction of some concepts that play an important role in the rest of the paper. We say that a strip P (from an element $(P, Q, n) \in \mathcal{R}(I)$, with P crossing L_s) is I-*transverse* if one can find finitely many elements $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{R}(I)$ such that

- every point in L_u which stays in $U \cup V$ under negative iteration is contained in one of the Q_α;
- every Q_α is I-transverse to P.

Strips P crossing L_s which are not I-transverse are said to be I-*critical*. The I-critical strips can be viewed as representing some "critical region" at the |I|-scale. One defines symmetrically I-transversality and I-criticality for horizontal strips Q crossing L_u. Elements $(P, Q, n) \in \mathcal{R}(I)$ such that both P and Q are I-critical are said to be I-*bicritical* and correspond to returns from the critical region to itself (in time n). Given a constant $\beta > 1$, one says that the parameter interval is β-*regular* if any I-bicritical (P, Q, n) 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. The value that we choose for $\beta > 1$ is announced in Section 5.6 and explained in Section 9.3. It depends only on the transverse Hausdorff dimensions d_s^0, d_u^0 and the eigenvalues of the periodic points p_s, p_u. It is easy to see that the starting interval I_0 is β-regular.

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

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

An easy and very natural property of the transversality relation, proved in Section 5.6, is that transversality is hereditary: let $P'_1 \subset P_1$, $Q'_0 \subset Q_0$, $I' \subset I$; if $Q_0 \pitchfork_I P_1$ holds, then $Q'_0 \pitchfork_{I'} P'_1$ also holds. A non-intuitive property of the transversality relation, but one which is useful at many places, is a partial converse called concavity (Section 6.3): with P_1, P'_1, Q_0, Q'_0, I, I′ as above, 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; 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; 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.

The most important result in Section 6 is a structure theorem for new rectangles in Section 6.5. One considers an element (P, Q, n) which belongs to $\mathcal{R}(I)$ but not to $\mathcal{R}(\widetilde{I})$, where $\widetilde{I}$ is the parameter interval containing I of the level immediately inferior (one says that $\widetilde{I}$ is the parent of I). Then there is a unique way to write (P, Q, n) as the result of a sequence of $k \geqslant 1$ parabolic compositions, possible in $\mathcal{R}(I)$ but not in $\mathcal{R}(\widetilde{I})$, of elements $(P_0, Q_0, n_0), \ldots, (P_k, Q_k, n_k) \in \mathcal{R}(\widetilde{I})$. This fundamental result is used with decisive effect at several places in Sections 7 and 9.

In Section 6.6 are proven several estimates on the width of strips, based on the structure theorem. As elements of $\mathcal{R}(I)$ are constructed using parabolic composition, one cannot hope for a uniform exponential estimate for the width of the strips in terms of the number n of iterations. However, we are able to prove a stretched exponential uniform estimate in Section 6.6.2. Another important result of Section 6.6 is that "thick" I-critical rectangles actually belong to some class $\mathcal{R}(I^*)$ for a parameter interval much bigger than I. Actually, because the definition of I-criticality supposes the knowledge of the full class $\mathcal{R}(I)$, which has not been constructed at this point, we introduce a substitute for this: a strip P (with $(P, Q, n) \in \mathcal{R}(I)$) is I-special if the thinnest $\widetilde{I}$-defined strip $\widetilde{P}$ (with $\widetilde{I}$ the parent of I) is $\widetilde{I}$-critical.

We complete in Section 7 the proof of the existence of the class $\mathcal{R}(I)$, for a candidate interval whose parent is β-regular. Actually, the long calculations are performed in Appendix A and we use in Section 7 the estimates derived from these calculations. We check that the uniform cone condition of the affine-like iterates in $\mathcal{R}(I)$ is satisfied, and that their distortion is bounded. After obtaining in Section 7.4 estimates for the class $\mathcal{R}(I_0)$ of the starting interval (which are simpler and better because no parabolic composition is involved), we prove in Section 7.5 a technical estimate related to parabolic composition. In Section 7.6, we deal with the relative speed of the strips when the parameter varies (the derivative of the quantity $\delta(Q, P')$ which comes up in the transversality relation); 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 speed estimates for all pairs of strips (actually, it is easy to see that such estimates do not exist); we have to restrict ourselves to I-special strips, a case which, fortunately, is sufficient for our purposes. In Section 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 bounded by a constant times the minimum), the result that we get (again only for I-special intervals) will allow us to argue as if it was.

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

Section 8 is a transition between the construction of the classes $\mathcal{R}(I)$ in Sections 5–7 and the heavy work of parameter selection in Section 9. We collect a number of properties of the transversality relation and the classes $\mathcal{R}(I)$ that were not needed before, but will be useful later. We also 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 investigate in particular (Section 8.2), for a given element $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$, the number of elements (P, Q, n) such that P is a non-simple child of $\widetilde{P}$; we show that, for every $\varepsilon > 0$, there are at most $\varepsilon^{-c'\eta}$ such non-simple children with width $|P|$ larger than $\varepsilon|\widetilde{P}|$. The constant η here is small and is the one occurring in 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 Section 8.3. Using Poincaré (or Dirichlet)-like series, we show that, for every $(P^*, Q^*, n^*) \in \mathcal{R}(I)$, every $\varepsilon > 0$, the number of elements $(P, Q, n) \in \mathcal{R}(I)$ with $P \subset P^*$, $|P| \geqslant \varepsilon|P^*|$ is at most $\varepsilon^{-d_s^*}$, with d_s^* very close to d_s^0. In Section 8.4, we transfer this information to parameter space, combining it with the result on relative speed of strips in Section 7.6: we show that, for any thin enough $\widetilde{I}$-critical strip Q^*, the proportion of candidates $I \subset \widetilde{I}$ such that Q^* is I-critical is at most $|\widetilde{I}|^{1-d_s^+}$, where again d_s^+ is close to d_s^0.

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 $\widetilde{I}$, most candidates $I \subset \widetilde{I}$ at the next level are β-regular. [It is a consequence of the structure theorem of Section 6 that all candidates are $\overline{\beta}$-regular, where $\overline{\beta} = \beta(1+\tau)^{-1}$ is very close to β; this allows 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 $|\widetilde{I}|^{\beta}$; it could happen a priori that for a regular parameter interval $\widetilde{I}$ we have many $\widetilde{I}$-bicritical $(P, Q, n) \in \mathcal{R}(I)$ with $|P|$ or $|Q|$ only slightly below the threshold $|\widetilde{I}|^{\beta}$ (and therefore above the next threshold $|I|^{\beta}$ for candidates $I \subset \widetilde{I}$); 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, and three sets of inequalities (SR1), (SR2), (SR3).

In the first part (Section 9.1), one controls the size of the critical locus, in two slightly different ways expressed by (SR1), (SR2); both amount to say that the "dimension" of the critical locus is not much larger than $d_s^0 + d_u^0 - 1$, but (SR1) is a direct "box-counting" estimate, while (SR2) is more subtle.

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 $|P| \geqslant x$ 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 Section 9.2 present a phase transition with respect to the width parameter x. Discussing this phase transition in Section 9.3 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 $\beta > 1$. We also show in Section 9.3 that a strongly regular parameter interval is regular.

In Section 9.4, we show that the starting interval is strongly regular. Then, in Section 9.5, we show that most candidates in a strongly regular parameter interval satisfy (SR1); The argument for parameter selection is based on the result of Section 8.4 mentioned above. The same is done in Section 9.6 for condition (SR2), but the proof is more complicated and involves (SR3).

In the rest of Section 9 we prove that, given a strongly regular parameter interval $\widetilde{I}$, most candidates $I \subset \widetilde{I}$ at the next level satisfy (SR3) (the proportion of failed candidates turns out to be not larger than $C|I|^{\tau^2}$; the same is true for (SR1) and (SR2)). The easy cases are dealt with in Section 9.7. In Section 9.8, we estimate the number of bicritical elements in $\mathcal{R}(I)$ that are not $\widetilde{I}$-defined ($\widetilde{I}$ is the strongly regular parent of I). This is a rather long but straightforward calculation based on the structure theorem of Section 6.5. Then, we are left with the most difficult case $I = I_\alpha = I_\omega$, x small (below the threshold of phase transition), but not too small (the estimate is trivial in this last case). In this case, and in this case only, we are forced to take into account that for an $\widetilde{I}$-bicritical element (P, Q, n) to be I-bicritical, the two events (of the "random" variable I) "P is I-critical" and "Q is I-critical" must occur simultaneously. The strategy (explained at the end of Section 9.7) to deal with this case is to divide the class of I-bicritical rectangles in a (large, but not too large) number of adequately defined subclasses. These subclasses are defined in Section 9.9, their number is estimated in Section 9.10 and the calculation which validates the strategy is performed in Sections 9.11 and 9.12. In Section 9.13, we explain which (very few) modifications have to be made when dealing with the condition $|Q| \geqslant x$ instead of $|P| \geqslant x$ (the setting is not symmetric at this level because the formulations of (SR3) when $d_s^0 \geqslant d_u^0$ and $d_s^0 \leqslant d_u^0$ are different). Finally, in Section 9.14 we conclude by defining precisely the exponents appearing in (SR1), (SR2), (SR3) which were so far only approximately defined. At this point, we know that all parameters in the starting interval $I_0 = [\varepsilon_0, 2\varepsilon_0]$, except for a subset of relative measure $\leqslant C\varepsilon_0^{\tau^2}$, are strongly regular.

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 $\mathcal{R}_+^\infty$, their union by $\widetilde{\mathcal{R}}_+^\infty$. We show in Section 10.5 that $\widetilde{\mathcal{R}}_+^\infty$ is a lamination by C^{1+Lip} curves with Lipschitz holonomy. In order to define a map on $\widetilde{\mathcal{R}}_+^\infty$ (which is not invariant under g), we introduce in Section 10.1 the concept of prime element in $\mathcal{R}$, i.e. one which cannot be written as the simple composition of shorter elements. The number of prime factors in a decomposition is controlled in Section 10.2. 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 $T^+(\omega) = \omega'$, $\widetilde{T}^+/\omega = g^n/\omega$. This defines a map T^+ from a subset $\mathcal{D}_+$ of $\mathcal{R}_+^\infty$ onto $\mathcal{R}_+^\infty$ which lifts to a map $\widetilde{T}^+$ from the union $\widetilde{\mathcal{D}}_+$ of curves in $\mathcal{D}_+$ to $\widetilde{\mathcal{R}}_+^\infty$.

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 $\mathcal{R}_+^\infty$ 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 Section 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 lamination $\widetilde{\mathcal{R}}_+^\infty$. This is proved in Section 10.10 (Theorem 4), but the proof is more difficult than in the classical case, because the transverse geometry of $\widetilde{\mathcal{R}}_+^\infty$ is complicated and requires some delicate handling.

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 $\widetilde{\mathcal{R}}_+^\infty - \widetilde{\mathcal{D}}_+$ where $\widetilde{T}^+$ 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 $\widetilde{T}^+$-invariant measure $\widetilde{\nu}$ 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 $\widetilde{\mathcal{R}}_+^\infty$ which was studied in Section 10. In the first part (Sections 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 $\mathcal{R}_+^\infty$ 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 smaller by a definite amount 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 repellors 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 repellor 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]. We also perform a number of estimates, both for simple and parabolic composition, which are used in Section 7.

In Appendix B, we prove a result related to Proposition 51 in Section 10.5, which states the transversally Lipschitz regularity of the lamination $\widetilde{\mathcal{R}}_+^\infty$.

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

We wish to thank Sylvain Crovisier for carefully reading a first version of this text and making many valuable suggestions to improve several conceptual parts of the paper, hopefully making it more amenable to read. We also wish to thank W. de Melo and M. Viana for fruitful conversations.

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

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

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 $R = \bigcup_{\mathcal{A}} R_a$. We choose the charts in order to have:

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

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

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

(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. Moreover, no rectangle R_a meets both the orbit of p_s and the orbit of p_u.

Let

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

The Markov partition provides a coding which is a topological conjugacy between the horseshoe K_g and the subshift of finite type of $\mathcal{A}^{\mathbf{Z}}$ 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 $W^s_{\mathrm{loc}}(p_s)$ is $\{x_{a_s} = 0\}$, the equation of the local unstable manifold $W^u_{\mathrm{loc}}(p_u)$ is $\{y_{a_u} = 0\}$.

We have written x_a (resp. y_a) for the coordinate in I^s_a (resp. I^u_a). We also choose the rectangles R_a in order to have, for some integer $N_0 \geqslant 2$:

(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 $L_u \subset \operatorname{int} R_{a_u}$. Taking the image under g^{N_0}, we get another lenticular region $L_s \subset \operatorname{int} R_{a_s}$. These regions are called *parabolic tongues*. See Figure 3.

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

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

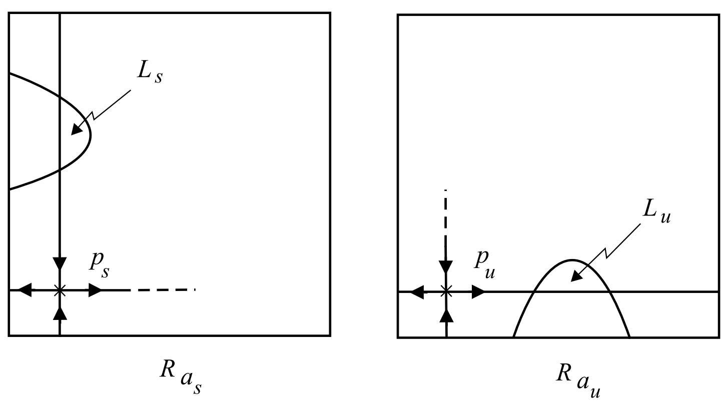

Fig. 3. — The parabolic tongues L_s, L_u

The maximal invariant set we are interested in is

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

Observe that $W^s_{\mathrm{loc}}(p_s)$ bounds the set K_g in R_{a_s} and $W^u_{\mathrm{loc}}(p_u)$ bounds the set K_g in R_{a_u}. The maximal invariant set Λ_g meets the boundary of $\widehat{R}$ (in fact, the boundary of L_s, L_u, and their images), but it is still locally maximal because it would also be the maximal invariant set in a slight enlargement of $\widehat{R}$.

We also define

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

$$(\mathbf{2.7})\qquad W^u(\Lambda_g, \widehat{R}) = \bigcap_{n\leqslant 0} g^{-n}(\widehat{R}).$$

The dynamics in $\widehat{R}$ are generated by

– the transition maps related to the Markov partition:

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

– 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 $I^s_{a_u} \times I^u_{a_u} \times I^s_{a_s} \times I^u_{a_s}$. Denote by π the projection from $I^s_{a_u} \times I^u_{a_u} \times I^s_{a_s} \times I^u_{a_s}$ onto $I^u_{a_u} \times I^s_{a_s}$. 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 singulary theory).

Denote by $\Gamma_0 \subset I^u_{a_u} \times I^s_{a_s}$ the smooth curve which is the image of the critical locus of this fold map. It divides $I^u_{a_u} \times I^s_{a_s}$ 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 $I^u_{a_u} \times I^s_{a_s}$ 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 $I^u_{a_u} \times I^s_{a_s}$.

Remark **2.** — The choice of θ is far from unique. One could for instance choose θ of the form

$$\theta(y_u, x_s) = \varepsilon_u y_u + \varepsilon_s \chi(x_s), \tag{2.8}$$

with ε_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

$$w^2 = \theta \circ \pi \tag{P4}$$

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

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

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

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 $G_+ \circ G_0 \circ G_-$:

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

with

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

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.

A simple explicit example for G,G_+,G_-,G_0,θ is given in [PY2].

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*

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

and whose image is a *horizontal strip*

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

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

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

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 $A_x, A_y, A_{xx}, B_x, B_y, \ldots$ for partial derivatives.

On the graph of F, we have

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

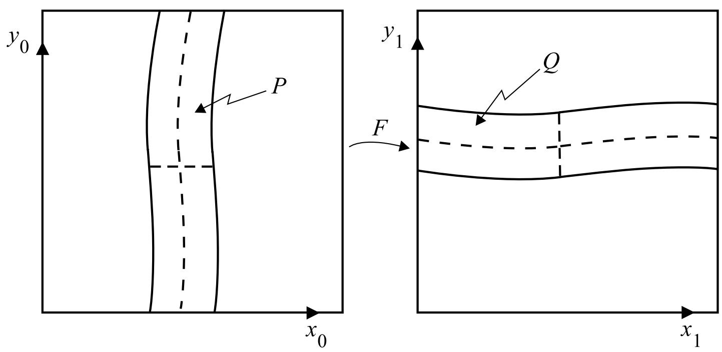

FIG. 4. — An affine-like map

which leads to

$$DF = A_x^{-1}\begin{pmatrix} 1 & -A_y \\ B_x & A_xB_y - A_yB_x \end{pmatrix}, \tag{3.3}$$

$$DF^{-1} = B_y^{-1}\begin{pmatrix} A_xB_y - A_yB_x & A_y \\ -B_x & 1 \end{pmatrix}, \tag{3.4}$$

$$\det DF = A_x^{-1}B_y. \tag{3.5}$$

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

3.2. *Cone condition and distortion.* — Let λ, u, $v > 0$ satisfy

$$1 < uv \leqslant \lambda^2. \tag{3.6}$$

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

This is readily seen to be equivalent to

(**AL′2**)
$$\lambda|A_x| + u|A_y| \leqslant 1,$$
$$\lambda|B_y| + v|B_x| \leqslant 1,$$

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

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

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

$$|P| := \max|A_x|, \tag{3.7}$$

and the width of the image Q by

$$|Q| := \max|B_y|. \tag{3.8}$$

Observe that we have $\max|A_x| \leqslant C \min|A_x|$, where the constant C only depends on D(F) and the lengths of the intervals I_0^u,I_1^s.The same estimate holds for B_y.

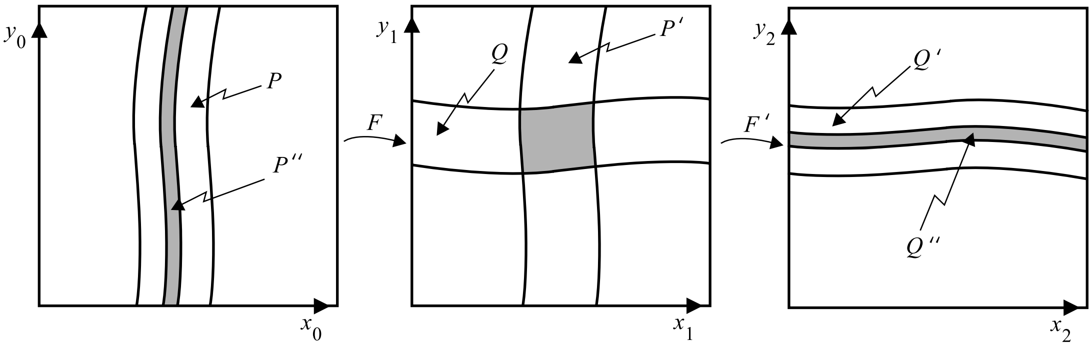

FIG. 5. — Simple composition

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 $F : P \to Q$ and $F' : P' \to Q'$ 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.

Let (A, B), (A′, B′), (A″, B″) be the implicit representations of F, F′, F″ respectively. Define

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

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

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

$$\begin{aligned} A''_y &= A_y + A'_y A_x B_y \Delta^{-1}, \\ B''_x &= B'_x + B_x A'_x B'_y \Delta^{-1}. \end{aligned} \tag{3.11}$$

From (3.10), we get

$$\begin{aligned} C^{-1} &\leqslant \frac{|P''|}{|P||P'|} \leqslant C, \\ C^{-1} &\leqslant \frac{|Q''|}{|Q||Q'|} \leqslant C, \end{aligned} \tag{3.12}$$

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. From formulas (A.12) to (A.21), one obtains the following estimate for the distortion:

$$D(F'') \leqslant \max\Big\{D(F) + C|Q|(D(F)+D(F')), D(F') + C|P'|(D(F)+D(F'))\Big\}, \tag{3.13}$$

where C depends only on u, v.

3.4. *Properties of the Markov partition.* — We choose charts for the Markov partition discussed in Section 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) $D(g_{\underline{a}}) \le D_0$.

Finally, the following simple property will be useful in Section 8.

(MP7) for any $(a, a') \in \mathcal{B}$, any $(x, y) \in P_{aa'}$, any $(x', y') \in Q_{aa'}$, we have

$$x_a^- + C^{-1} < x < x_a^+ - C^{-1},$$
$$y_{a'}^- + C^{-1} < y' < y_{a'}^+ - C^{-1},$$

where $I_a^s = [x_a^-, x_a^+]$, $I_{a'}^u = [y_{a'}^-, y_{a'}^+]$ and C is an appropriate large constant.

3.5. *Parabolic composition.* — Let G be the folding map of Section 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.

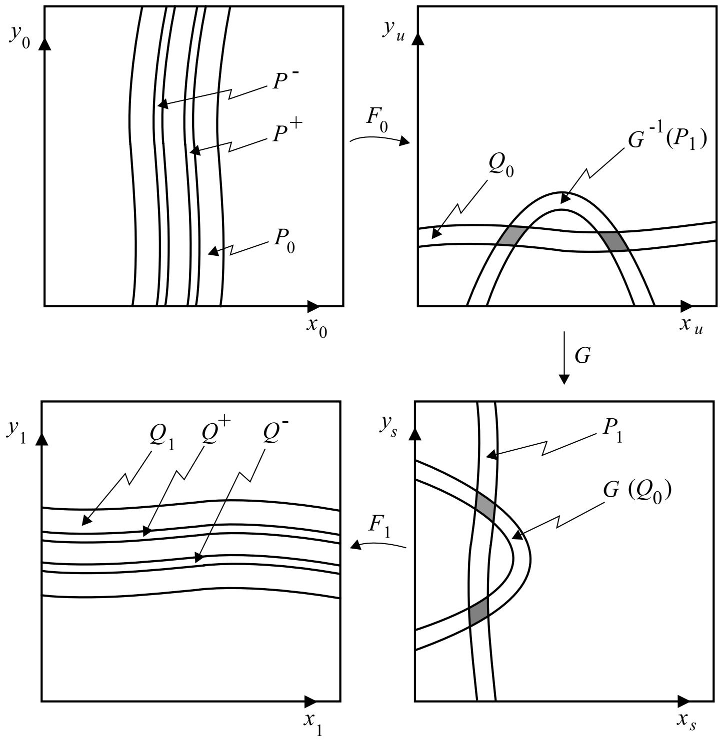

FIG. 6. — Parabolic composition

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

$$\textbf{(PC1)}\qquad |A_{1,y}| < b, \qquad |A_{1,yy}| < b,$$
$$|B_{0,x}| < b, \qquad |B_{0,xx}| < b,$$

with $b \ll 1$. In the system

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

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

$$x_u = X(w, y_0). \tag{3.15}$$

Similarly, in the system

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

we eliminate x_s, solve for y_s to get

$$y_s = Y(w, x_1). \tag{3.17}$$

The next step is to define

$$C(w, y_0, x_1) := w^2 - \theta\Big(B_0(y_0, X(w, y_0)), A_1(Y(w, x_1), x_1)\Big). \tag{3.18}$$

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

$$\gamma_0 = \Big\{y_u = B_0(y_0^*, X(w, y_0^*))\Big\}; \tag{3.19}$$

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

$$\gamma_1 = \Big\{x_s = A_1(Y(w, x_1^*), x_1^*)\Big\}. \tag{3.20}$$

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

$$|C_w - 2w| \ll 1, \tag{3.21}$$

$$|C_{ww} - 2| \ll 1. \tag{3.22}$$

Therefore, for fixed values of y_0 and x_1, C has a unique minimum as a function of w; we denote by $\overline{C}(y_0, x_1)$ the corresponding minimum value. We have $\overline{C}(y_0^*, x_1^*) > 0$ (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 $\overline{C}(y_0, x_1) < 0$ everywhere on $I_0^u \times I_1^s$. Setting

$$\delta = \delta(Q_0, P_1) = \min_{y_0, x_1} -\overline{C}(y_0, x_1) \tag{3.23}$$

we actually want to have

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

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

$$w = W^{\pm}(y_0, x_1) \tag{3.24}$$

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

$$A^{\pm}(y_0, x_1) := A_0\Big(y_0, X(W^{\pm}(y_0, x_1), y_0)\Big), \tag{3.25}$$

$$B^{\pm}(y_0, x_1) := B_1\Big(Y(W^{\pm}(y_0, x_1), x_1), x_1\Big). \tag{3.26}$$

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^+ (resp. F^-) and by Q^+ (resp. Q^-) the image of F^+ (resp. 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 (see in particular (A.42)). The partial derivative C_w in these formulas is of order $\delta^{\frac{1}{2}}$ at the points under consideration (see (A.84) in Appendix A). From this one obtains the following estimate for the widths:

$$C^{-1} \leqslant \frac{|P^{\pm}|}{|P_0||P_1|\delta^{-\frac{1}{2}}} \leqslant C, \tag{3.27}$$

$$C^{-1} \leqslant \frac{|Q^{\pm}|}{|Q_0||Q_1|\delta^{-\frac{1}{2}}} \leqslant C, \tag{3.28}$$

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 $F^{\pm}$ (see also Appendix A.4): assuming that b is small enough (in terms of the partial derivatives of first order of X_u, Y_s, θ), we have

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

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.

Let us observe that, while conditions (PC1), (PC2) are *necessary* in order to consider parabolic composition, they will not be *sufficient* for our construction: in Section 5, the requirement for parabolic composition will be much more restrictive than (PC2).

3.6. *Simple properties of* $\delta(Q_0, P_1)$ *and related quantities.*

3.6.1. The setting in the same than in the last Subsection, with maps F_0, F_1 satisfying the hypotheses as in that Subsection. We have defined

$$\delta(Q_0, P_1) = \min_{y_0} \min_{x_1} -\overline{C}(y_0, x_1). \tag{3.30}$$

In the following sections, we also need to consider

$$\delta_L(Q_0, P_1) := \max_{y_0} \min_{x_1} -\overline{C}(y_0, x_1), \tag{3.31}$$

$$\delta_R(Q_0, P_1) := \min_{y_0} \max_{x_1} -\overline{C}(y_0, x_1), \tag{3.32}$$

$$\delta_{LR}(Q_0, P_1) := \max_{y_0} \max_{x_1} -\overline{C}(y_0, x_1). \tag{3.33}$$

All together, δ, δ_L, δ_R, δ_{LR} are the values of $-\overline{C}$ at the four corners of the rectangle $I_0^u \times I_1^s$ of definition of $\overline{C}$. From the formula (A.78) of Appendix A, we have, for any t in I

$$C^{-1}|P_1| \leqslant |\overline{C}_x| \leqslant C|P_1|, \tag{3.34}$$

$$C^{-1}|Q_0| \leqslant |\overline{C}_y| \leqslant C|Q_0|. \tag{3.35}$$

This gives

$$C^{-1}|Q_0| \leqslant \delta_L(Q_0, P_1) - \delta(Q_0, P_1) \leqslant C|Q_0|, \tag{3.36}$$

$$C^{-1}|P_1| \leqslant \delta_R(Q_0, P_1) - \delta(Q_0, P_1) \leqslant C|P_1|, \tag{3.37}$$

$$C^{-1}|Q_0| \leqslant \delta_{LR}(Q_0, P_1) - \delta_R(Q_0, P_1) \leqslant C|Q_0|, \tag{3.38}$$

$$C^{-1}|P_1| \leqslant \delta_{LR}(Q_0, P_1) - \delta_L(Q_0, P_1) \leqslant C|P_1|. \tag{3.39}$$

Let F'_0 be another affine-like map, implicitly represented by (A_0, B_0), satisfying the same hypotheses than F_0. Let C' be the function constructed from F'_0 and F_1 as C was from F_0 and F_1.

3.6.2. We first assume that the image Q'_0 of F'_0 is *disjoint* from the image Q_0 of F_0.

Assume for instance that Q'_0 is above Q_0. For any $y_0^* \in I_0^u, y_0'^* \in I_0'^u$, the curve $G_- \circ F'_0(\{y'_0 = y_0'^*\})$ is above the curve $G_- \circ F_0(\{y_0 = y_0^*\})$. This means that we have, for any w

$$B_0(y_0^*, X(w, y_0^*)) \leqslant B'_0(y_0'^*, X'(w, y_0'^*)). \tag{3.40}$$

As θ is monotone in the first variable, we will have, for all $w, x_1, y_0^*, y_0'^*$

$$C(w, y_0^*, x_1) < C'(w, y_0'^*, x_1), \tag{3.41}$$

if $\theta_y > 0$, and the opposite inequality if $\theta_y < 0$. In the first case, we will have

$$\delta(Q_0, P_1) > \delta_L(Q'_0, P_1), \tag{3.42}$$

$$\delta_R(Q_0, P_1) > \delta_{LR}(Q'_0, P_1). \tag{3.43}$$

In the second case, the same is true with Q_0, Q'_0 exchanged.

3.6.3. Assume now that the image Q'_0 of F'_0 is *contained* in Q_0.

Let $y_0^{\pm}$ be the endpoints of I_0^u. For any $y_0'^*$, the curve $G_- \circ F'_0(\{y'_0 = y_0'^*\})$ is between the curves $G_- \circ F_0(\{y_0 = y_0^{\pm}\})$. This means that we have, for any w

$$B_0(y_0^-, X(w, y_0^-)) \leqslant B'_0(y_0'^*, X'(w, y_0'^*)) \leqslant B_0(y_0^+, X(w, y_0^+)) \tag{3.44}$$

(assuming $B_{0,y} > 0$; otherwise, exchange y_0^+ and y_0^-).

As θ is monotone in the first variable, we have, for all w, $x_1, y_0'^*$

$$\min_{y_0} C(w, y_0, x_1) \leqslant C'(w, y_0'^*, x_1) \leqslant \max_{y_0} C(w, y_0, x_1). \tag{3.45}$$

This now gives

$$\delta(Q_0, P_1) \leqslant \delta(Q'_0, P_1), \tag{3.46}$$

$$\delta_R(Q_0, P_1) \leqslant \delta_R(Q'_0, P_1), \tag{3.47}$$

$$\delta_L(Q_0, P_1) \geqslant \delta_L(Q'_0, P_1), \tag{3.48}$$

$$\delta_{LR}(Q_0, P_1) \geqslant \delta_{LR}(Q'_0, P_1). \tag{3.49}$$

3.6.4. The setting is the same than in Section 3.6.2 and we will obtain slightly stronger estimates (which will be useful in Section 8) under a slightly stronger hypothesis. We not only assume that the image Q'_0 of F'_0 is contained in Q_0, but that one has, for any $(x_0, y_0) \in F_0^{-1}(Q'_0)$

$$y_0^- + C^{-1} < y_0 < y_0^+ - C^{-1}, \tag{3.50}$$

for some fixed large constant C. We will now obtain, instead of (3.44)

$$\begin{aligned} B_0(y_0^- + C^{-1}, X(w, y_0^- + C^{-1})) &\leqslant B'_0(y_0'^*, X'(w, y_0'^*)) \\ &\leqslant B_0(y_0^+ - C^{-1}, X(w, y_0^+ - C^{-1})) \end{aligned} \tag{3.51}$$

or the reverse inequalities. This gives

$$\begin{aligned} B_0(y_0^-, X(w, y_0^-)) + C^{-1}|Q_0| &\leqslant B'_0(y_0'^*, X'(w, y_0'^*)) \\ &\leqslant B_0(y_0^+, X(w, y_0^+)) - C^{-1}|Q_0| \end{aligned} \tag{3.52}$$

(if $B_{0,y} > 0$; otherwise, exchange y_0^+ and y_0^-). As the partial derivative θ_y is bounded away from 0, we get also, for all w, $x_1, y_0'^*$

$$\min_{y_0} C(w, y_0, x_1) + C^{-1}|Q_0| \leqslant C'(w, y_0', x_1) \leqslant \max_{y_0} C(w, y_0, x_1) - C^{-1}|Q_0| \tag{3.53}$$

which finally gives

$$\delta(Q_0, P_1) + C^{-1}|Q_0| \leqslant \delta(Q_0', P_1), \tag{3.54}$$

$$\delta_R(Q_0, P_1) + C^{-1}|Q_0| \leqslant \delta_R(Q_0', P_1), \tag{3.55}$$

$$\delta_L(Q_0, P_1) - C^{-1}|Q_0| \geqslant \delta_L(Q_0', P_1), \tag{3.56}$$

$$\delta_{LR}(Q_0, P_1) - C^{-1}|Q_0| \geqslant \delta_{LR}(Q_0', P_1). \tag{3.57}$$

3.6.5. All the discussion in Sections 3.6.2, 3.6.3, 3.6.4 have an obvious symmetric counterpart exchanging Q's and P's.

4. Structure of parameter space

4.1. *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 Section 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 Section 2.3 choose a function θ which depends smoothly on t.

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

$$\frac{\partial}{\partial t}\theta(0, 0, t)\,|_{t=0} > 0. \tag{4.1}$$

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

$$\theta(0, 0, t) \equiv t, \qquad t \in (-t_0, t_0). \tag{4.2}$$

4.2. *Some important constants.* — The constants λ, u, v, D_0, b of Sections 3.4 and 3.5 depend only on the initial diffeomorphism f provided $\mathcal{U}$ is small enough and are now fixed.

Throughout the rest of the paper, we will use three main constants ε_0, η, τ which satisfy

$$0 < \varepsilon_0 \ll \eta \ll \tau \ll 1. \tag{4.3}$$

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 Section 3.5, roughly speaking we will ask that

$$\delta \geqslant (|P_1| + |Q_0|)^{1-\eta}. \tag{4.4}$$

As the widths $|P_1|$ and $|Q_0|$ are not larger than ε_0 which is much smaller than η, this requirement is much more restrictive than (PC2) and will allow to control distortion.
– τ relates the successive scales of the parameter intervals we will consider through the formula $\varepsilon_{k+1} = \varepsilon_k^{1+\tau}$.

Another important number β appears in the definition of regularity in Section 5, which controls the recurrence of the "critical locus". This number will be chosen explicitly in Section 9.3 (see also Section 5.6) in terms of d_s^0, d_u^0 and the eigenvalues of f at the periodic points p_s, p_u; the condition (H4) involving d_s^0, d_u^0 in Section 1 is required because we must have $\beta > 1$.

Finally, we will use the generic letter C (with indices or other decorations) for various constants which depend on the initial diffeomorphism f, but not on ε_0, η, τ.

4.3. *Parameter intervals.* — The starting parameter interval will be

$$I_0 := [\varepsilon_0, 2\varepsilon_0], \tag{4.5}$$

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

$$\varepsilon_{k+1} = \varepsilon_k^{1+\tau}. \tag{4.6}$$

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

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

$$\sum_{k\geqslant 0} \alpha_k \ll 1, \tag{4.7}$$

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 **1.** — *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* $t \in \mathrm{I}$, *the restriction of* g_t^n *to* P_t *is an affine-like map onto* Q_t, *i.e. property* (AL1) *of Section* 3.1 *holds;*
– *for each* $t \in \mathrm{I}$, *each* $m \in [0, n]$, *we have* $g_t^m(\mathrm{P}_t) \subset \widehat{\mathrm{R}}$.

Example **1.**

1. For $n = 0$, the I-persistent affine-like iterates are the $(\mathrm{R}_a, \mathrm{R}_a, 0)$, $a \in \mathcal{A}$.
2. For $n = 1$, the I-persistent affine-like iterates are the $(\mathrm{P}_{aa'}, \mathrm{Q}_{aa'}, 1)$, $(a, a') \in \mathcal{B}$.
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 Section 3.4 defined an I-persistent affine-like iterate $(\mathrm{P}_{\underline{a}}, \mathrm{Q}_{\underline{a}}, n)$.

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

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

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

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

Remark **3.** — Throughout the paper, except in Section 9 (where we break the symmetry assuming $d_s^0 \geqslant d_u^0$), 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 $t \in \mathrm{I}$, P is a connected component of $\mathrm{R} \cap g_t^{-n}(\mathrm{R})$ and also of $\bigcap_{0 \leqslant m \leqslant n} g_t^{-m}(\widehat{\mathrm{R}})$.

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

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

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

5.2. *The classes* $\mathcal{R}(\mathrm{I})$*: 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 construct, for every candidate parameter interval I (see Section 4.4: it means that $I = I_0$ or that the parent interval of I is regular, a notion that will be defined in Section 5.6) a subset $\mathcal{R}(I)$ of the set of all I-persistent affine-like iterates. All elements of $\mathcal{R}(I)$ with $n > 1$ will be obtained, from the very definition of $\mathcal{R}(I)$, 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 Section 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.

The precise requirements on $\mathcal{R}(I)$ are the properties (R1)–(R6) formulated in the next subsection and property (R7) formulated in Section 5.4.

While it is clear that at most one class $\mathcal{R}(I)$ can satisfy (R1)–(R7), it is by no means obvious that such a class exists. Actually, the proof of this fact, stated at the end of Section 5.4, will only be completed at the end of Section 7.

The proof of the existence of $\mathcal{R}(I)$ is based on a double induction:

– an induction on the level of the parameter interval I, starting with $I_0 = [\varepsilon_0, 2\varepsilon_0]$ at level 0 (see Section 5.5 for this first step).
– for a given candidate parameter interval I, an induction on the length n of the I-persistent affine-like iterates under consideration.

In this induction scheme, all properties required for $\mathcal{R}(I)$ are proved simultaneously. Actually, several other properties (coherence, concavity, ... see Section 6) of the class $\mathcal{R}(I)$ are needed in the induction process; these properties are defined and proved (inductively!) in Section 6. Of particular importance is a structure theorem in Section 6.5.

5.3. *Defining properties for the special class of affine-like iterates* $\mathcal{R}(I)$. — Let I be a candidate parameter interval of some level.

The class $\mathcal{R}(I)$ of I-persistent affine-like iterates that we want to construct should satisfy the following properties (R1)–(R7).

(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 1 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 Section 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 $1 < uv \leqslant \lambda^2$), and have distortion bounded by D_0. Let $u_0 = \frac{u}{(uv)^{1/4}}$, $v_0 = \frac{v}{(uv)^{1/4}}$.

(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 Section 3.3 that the simple composition defined by

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

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 $(P_{\underline{a}}, Q_{\underline{a}}, n)$ 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.

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

$$\begin{aligned} &C^{-1}\varepsilon_0 \leqslant |P_s| \leqslant C\varepsilon_0, \\ &C^{-1}\varepsilon_0 \leqslant |Q_u| \leqslant C\varepsilon_0. \end{aligned} \tag{5.2}$$

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

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

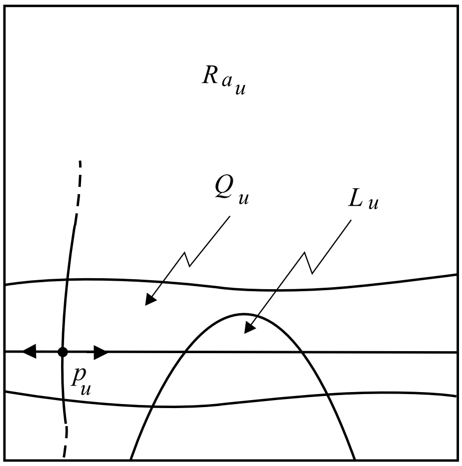

FIG. 7. — The special rectangle Q_u

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

$$|A_y| \leqslant C\varepsilon_0, \qquad |A_{yy}| \leqslant C\varepsilon_0.$$

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

$$|B_x| \leqslant C\varepsilon_0, \qquad |B_{xx}| \leqslant C\varepsilon_0.$$

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 Section 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 Section 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, Section 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), $(\widetilde{P}, \widetilde{Q}, \widetilde{n}) \in \mathcal{R}(I)$ with $P \subset \widetilde{P}$, $n > \widetilde{n}$. When there is no $(\widehat{P}, \widehat{Q}, \widehat{n}) \in \mathcal{R}(I)$ with $P \subset \widehat{P} \subset \widetilde{P}$ and $n > \widehat{n} > \widetilde{n}$, we say that P is a *child* of $\widetilde{P}$ and that $\widetilde{P}$ is the *parent* of P.

When $n = \widetilde{n} + 1$, we say that P is a *simple* child of $\widetilde{P}$. The image $g^{\widetilde{n}}(P)$ is equal to the intersection of $\widetilde{Q}$ with some $P_{a,a'}$ and (P, Q, n) is the simple composition of $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$ with $(P_{a,a'}, Q_{a,a'}, 1)$.

When $n > \widetilde{n} + 1$, we say that P is a *non-simple* child of $\widetilde{P}$. The image $g^{\widetilde{n}}(P)$ is contained in $L_u \cap \widetilde{Q}$. We will prove in Section 6.2 that (P, Q, n) is obtained by parabolic composition of $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$ with some (P_1, Q_1, n_1)

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

$$(P, Q, n) = (P_0, Q_0, n_0) * (P_1, Q_1, n_1). \tag{5.3}$$

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

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

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 Section 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 Section 3.5, together with

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

In Section 3.5, we were asking for δ to be much larger than $|P_0|$ and $|Q_1|$. Recall from Section 3.6 the definitions

$$\delta_L(Q_0, P_1) := \max_{y_0} \min_{x_1} -\overline{C}(y_0, x_1), \tag{5.6}$$

$$\delta_R(Q_0, P_1) := \min_{y_0} \max_{x_1} -\overline{C}(y_0, x_1), \tag{5.7}$$

$$\delta_{LR}(Q_0, P_1) := \max_{y_0} \max_{x_1} -\overline{C}(y_0, x_1). \tag{5.8}$$

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

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

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

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

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

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

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

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* $\widetilde{I} \supset I$, *elements* $(\widetilde{P}_0, \widetilde{Q}_0, \widetilde{n}_0)$, $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1) \in \mathcal{R}(\widetilde{I})$ *with* $\widetilde{P}_1 \supset P_1$, $\widetilde{Q}_0 \supset Q_0$ *such that* $\widetilde{Q}_0 \overline{\pitchfork}_{\widetilde{I}} \widetilde{P}_1$.

Remark **4.**

1. Taking $\widetilde{\mathrm{I}} = \mathrm{I}$, $\widetilde{\mathrm{P}}_0 = \mathrm{P}_0$, $\widetilde{\mathrm{Q}}_1 = \mathrm{Q}_1$, it is obvious that if $\mathrm{Q}_0 \,\overline{\pitchfork}_{\mathrm{I}}\, \mathrm{P}_1$, then $\mathrm{Q}_0 \pitchfork_{\mathrm{I}} \mathrm{P}_1$.
2. In view of our inductive procedure, all $(\widetilde{\mathrm{P}}_0, \widetilde{\mathrm{Q}}_0, \widetilde{n}_0)$, $(\widetilde{\mathrm{P}}_1, \widetilde{\mathrm{Q}}_1, \widetilde{n}_1)$ which have to be considered have been constructed before $(\mathrm{P}_0, \mathrm{Q}_0, n_0)$, $(\mathrm{P}_1, \mathrm{Q}_1, n_1)$.
3. As mentioned before, the definition of the transversality relation is quite involved. One reason for this is that we wish to obtain several properties (heredity, concavity, ...) that will be proven in the next subsections. 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 (Section 3.5), because they involve only one value of the parameter. The next property takes care of this problem.

(R7) If $(\mathrm{P}_0, \mathrm{Q}_0, n_0)$, $(\mathrm{P}_1, \mathrm{Q}_1, n_1) \in \mathcal{R}(\mathrm{I})$ satisfy $\mathrm{Q}_0 \subset \mathrm{Q}_u$, $\mathrm{P}_1 \subset \mathrm{P}_s$ and $\mathrm{Q}_0 \pitchfork_{\mathrm{I}} \mathrm{P}_1$ holds, then, for all $t \in \mathrm{I}$, we have

$$\delta(\mathrm{Q}_0, \mathrm{P}_1) \geqslant \mathrm{C}^{-1}\Big(|\mathrm{P}_1|^{1-\eta} + |\mathrm{Q}_0|^{1-\eta}\Big).$$

Now that properties (R1)–(R7) have been introduced, we can state the

Theorem **1.** — *For each candidate parameter interval* I, *there exists exactly one class* $\mathcal{R}(\mathrm{I})$ *of I-persistent affine-like iterates which satisfies* (R1)–(R7).

As mentioned before, uniqueness is clear. The proof of existence will only be completed at the end of Section 7.

5.5. *The class* $\mathcal{R}(\mathrm{I}_0)$. — We claim that the class $\mathcal{R}(\mathrm{I}_0)$ is exactly formed by the iterates associated to the horseshoe K considered in Section 3.4, i.e. the elements $(\mathrm{P}_{\underline{a}}, \mathrm{Q}_{\underline{a}}, n)$ with $\underline{a} = (a_0, \dots, a_n)$ a word with transitions in $\mathcal{B}$.

Indeed, such elements must belong to $\mathcal{R}(\mathrm{I}_0)$ by (R1). They satisfy (R2), (R3) and (R6). We show below that, for $(\mathrm{P}_0, \mathrm{Q}_0, n_0)$,$(\mathrm{P}_1, \mathrm{Q}_1, n_1)$ of this form with $\mathrm{Q}_0 \subset \mathrm{Q}_u$,$\mathrm{P}_1 \subset \mathrm{P}_s$, the transversality relation $\mathrm{Q}_0 \pitchfork_{\mathrm{I}} \mathrm{P}_1$ is never satisfied. Then it follows that (R5), (R7) are vacuously satisfied. Finally, property (R4) will be proved in Proposition 16 of Section 7.4.

Let $(\mathrm{P}_0, \mathrm{Q}_0, n_0)$,$(\mathrm{P}_1, \mathrm{Q}_1, n_1)$ as above (with $\mathrm{Q}_0 \subset \mathrm{Q}_u$,$\mathrm{P}_1 \subset \mathrm{P}_s$). The notations are those of Sections 3.5 and 5.4. Assume for instance that in R_{a_u} the foliation $\mathrm{W}^u_{\mathrm{loc}}(\mathrm{K})$ is contained in $\{y_{a_u} \geqslant 0\}$, and that in R_{a_s} the foliation $\mathrm{W}^s_{\mathrm{loc}}(\mathrm{K})$ is contained in $\{x_{a_s} \geqslant 0\}$. This is equivalent to say that θ is monotone increasing in both variables.

There exists y_0^* such that the image under $\mathrm{G}_- \circ g^{n_0}$ of the horizontal segment $\{y = y_0^*\} \cap \mathrm{P}_0$ lies in $\{y_u > 0\}$. Similarly, there exists x_1^* such that the image under $\mathrm{G}_+^{-1} \circ g^{-n_1}$ of the vertical segment $\{x = x_1^*\} \cap \mathrm{Q}_1$ lies in $\{x_s > 0\}$. Then, from formulas (3.18)–(3.20) of Section 3.5 and the monotonicity of θ, we have, for all $t \in \mathrm{I}_0$ and w

$$\mathrm{C}(w, y_0^*, x_1^*) \geqslant -\theta(0, 0, t). \tag{5.9}$$

As we have $\theta(0,0,t) = t$ by the normalization of Section 4.1, we obtain, for $t = \varepsilon_0$:

$$(\mathbf{5.10}) \qquad -\overline{C}(y_0^*, x_1^*) \leqslant |I_0|,$$

and therefore, for the same value of t:

$$(\mathbf{5.11}) \qquad \delta(Q_0, P_1) \leqslant |I_0|.$$

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

$$(\mathbf{5.12}) \qquad \delta(Q_0, P_1) \leqslant 2|I_0|.$$

If condition (T1) in Section 5.4 is satisfied, we have from (3.36)–(3.39) that at $t = \varepsilon_0$ either $|Q_0|$ or $|P_1|$ (or both) is at least of the order of ε_0. But then Proposition 16 in Section 7.4 guarantees that the same width stays at least of the order of ε_0 for all $t \in I_0$. Then either (T2) or (T3) will not be satisfied.

We have thus shown that the transversality relation is never satisfied in $\mathcal{R}(I_0)$. This completes the initial step in the proof of the theorem in Section 5.4 (except for the proof of Proposition 16 in Section 7.4).

5.6. *Criticality, bicriticality, and the regularity property.* — We introduce some terminology and some concepts related to the transversality relation. This includes the regularity property mentioned in the selection process of Section 4.4.

5.6.1. The following obvious but fundamental property, which we may sum up by saying that transversality is hereditary, was forced into the definition of the transversality relation.

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

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

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

Corollary **2.** — *Let* $\widetilde{I} \supset I$ *be parameter intervals. Then* $\mathcal{R}(\widetilde{I})$ *is contained in* $\mathcal{R}(I)$.

Remark **5.** — This is a slight abus de langage of no consequence: properly speaking, we mean that the restriction to I of any $(P, Q, n) \in \mathcal{R}(\widetilde{I})$ belongs to $\mathcal{R}(I)$.

Proof. — This is a consequence, by induction of the length n, of Axiom (R6) and Corollary 1. □

The following result is also an easy consequence of the definition of the transversality relation.

Proposition **3.** — *Let* $\mathrm{I} \subset \widehat{\mathrm{I}}$ *be parameter intervals, and let* $(\mathrm{P}_0, \mathrm{Q}_0, n_0)$, $(\mathrm{P}_1, \mathrm{Q}_1, n_1)$ *be elements of* $\mathcal{R}(\mathrm{I})$ *such that* $\mathrm{Q}_0 \subset \mathrm{Q}_u$ *and* $\mathrm{P}_1 \subset \mathrm{P}_s$. *Assume that* $\mathrm{Q}_0 \pitchfork_{\mathrm{I}} \mathrm{P}_1$ *holds but* $\mathrm{Q}_0 \pitchfork_{\widehat{\mathrm{I}}} \mathrm{P}_1$ *does not hold. Then there exists* $t \in \widehat{\mathrm{I}}$ *such that* $\delta_{\mathrm{LR}}(\mathrm{Q}_0, \mathrm{P}_1) < 2|\widehat{\mathrm{I}}|$.

Proof. — Let $\widetilde{\mathrm{I}} \supset \mathrm{I}$, and let elements $(\widetilde{\mathrm{P}}_0, \widetilde{\mathrm{Q}}_0, \widetilde{n}_0)$, $(\widetilde{\mathrm{P}}_1, \widetilde{\mathrm{Q}}_1, \widetilde{n}_1) \in \mathcal{R}(\widetilde{\mathrm{I}})$ with $\widetilde{\mathrm{P}}_1 \supset \mathrm{P}_1$, $\widetilde{\mathrm{Q}}_0 \supset \mathrm{Q}_0$ such that $\widetilde{\mathrm{Q}}_0 \overline{\pitchfork}_{\widetilde{\mathrm{I}}} \widetilde{\mathrm{P}}_1$ holds. As $\mathrm{Q}_0 \pitchfork_{\widehat{\mathrm{I}}} \mathrm{P}_1$ does not hold, $\widetilde{\mathrm{I}}$ is strictly contained in $\widehat{\mathrm{I}}$ and $\widetilde{\mathrm{Q}}_0 \overline{\pitchfork}_{\widehat{\mathrm{I}}} \widetilde{\mathrm{P}}_1$ does not hold. But conditions (T2), (T3) for $\widetilde{\mathrm{Q}}_0 \overline{\pitchfork}_{\widetilde{\mathrm{I}}} \widetilde{\mathrm{P}}_1$ imply the same for $\widetilde{\mathrm{Q}}_0 \overline{\pitchfork}_{\widehat{\mathrm{I}}} \widetilde{\mathrm{P}}_1$. Hence (T1) for $\widetilde{\mathrm{Q}}_0 \overline{\pitchfork}_{\widehat{\mathrm{I}}} \widetilde{\mathrm{P}}_1$ does not hold and there exists $t \in \widehat{\mathrm{I}}$ such that $2|\widehat{\mathrm{I}}| > \delta_{\mathrm{LR}}(\widetilde{\mathrm{Q}}_0, \widetilde{\mathrm{P}}_1) \geqslant \delta_{\mathrm{LR}}(\mathrm{Q}_0, \mathrm{P}_1)$. □

5.6.2. Let I be a (candidate) parameter interval, and let $(\mathrm{P}_0, \mathrm{Q}_0, n_0)$, $(\mathrm{P}_1, \mathrm{Q}_1, n_1)$ be elements of $\mathcal{R}(\mathrm{I})$ such that $\mathrm{Q}_0 \subset \mathrm{Q}_u$ and $\mathrm{P}_1 \subset \mathrm{P}_s$. When Q_0 and P_1 are not I-transverse, we say that:

– Q_0 and P_1 are I-*separated* if $\mathrm{G}_t(\mathrm{Q}_0) \cap \mathrm{P}_1 \cap \Lambda = \emptyset$ for all $t \in \mathrm{I}$; this happens in particular when $\delta_{\mathrm{LR}}(\mathrm{Q}_0, \mathrm{P}_1) < 0$ for all $t \in \mathrm{I}$;
– Q_0 and P_1 are I-*critically related* otherwise.

5.6.3. We assume in this subsection that I is a parameter interval for which the class $\mathcal{R}(\mathrm{I})$ has been fully constructed with properties (R1)–(R7).

Let $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}(\mathrm{I})$. An I-*decomposition* of P is a finite family $(\mathrm{P}_\alpha, \mathrm{Q}_\alpha, n_\alpha)$ of elements of $\mathcal{R}(\mathrm{I})$ such that the P'_αs are disjoint, strictly contained in P and satisfy, for all $t \in \mathrm{I}$

$$\mathrm{W}^s(\Lambda, \widehat{\mathrm{R}}) \cap \mathrm{P} = \bigsqcup_\alpha (\mathrm{W}^s(\Lambda, \widehat{\mathrm{R}}) \cap \mathrm{P}_\alpha), \tag{5.13}$$

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

Remark **6.** — We will see in Section 8 that any P has only finitely many children.

Let $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}(\mathrm{I})$. We say that Q is I-*transverse* if either $\mathrm{Q} \cap \mathrm{Q}_u = \emptyset$ or $\mathrm{Q} \subset \mathrm{Q}_u$ and there exists an I-*decomposition* $(\mathrm{P}_\alpha, \mathrm{Q}_\alpha, n_\alpha)$ 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 $\mathrm{Q} \supset \mathrm{Q}_u$.

Let $(P, Q, n), (P', Q', n') \in \mathcal{R}(I)$ satisfy $Q' \subset Q$. As transversality is hereditary (Proposition 2), if Q is I-transverse, then Q' is also I-transverse.

All these notions are also symmetrically defined exchanging P's and Q's,future and past.

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 thought as a piece of the dynamics corresponding to a return of the "critical region" to itself.

Definition **2.** — *Let* $\beta > 1$. *We say that the candidate 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* $t \in I$*:*

$$|P| < |I|^{\beta}, \qquad |Q| < |I|^{\beta}. \tag{5.14}$$

5.6.4. For $i = s, u$, let $\lambda(p_i)$ (resp. $\mu(p_i)$) be the stable (resp. unstable) eigenvalue of the periodic point p_i. Define

$$\omega_s = -\frac{\log|\lambda(p_s)|}{\log|\mu(p_s)|}, \qquad \omega_u = -\frac{\log|\mu(p_u)|}{\log|\lambda(p_u)|}. \tag{5.15}$$

As we have

$$C^{-1}\varepsilon_0 \leqslant |P_s| \leqslant C\varepsilon_0, \qquad C^{-1}\varepsilon_0 \leqslant |Q_u| \leqslant C\varepsilon_0, \tag{5.16}$$

we will have

$$C^{-1}\varepsilon_0^{\omega_s} \leqslant |Q_s| \leqslant C\varepsilon_0^{\omega_s}, \qquad C^{-1}\varepsilon_0^{\omega_u} \leqslant |P_u| \leqslant C\varepsilon_0^{\omega_u}. \tag{5.17}$$

Remark **7.** — The ratios ω_s, ω_u are smooth functions of the parameter, therefore the relative variation of the quantities $\varepsilon_0^{\omega_s}, \varepsilon_0^{\omega_u}$ in I_0 is negligible.

Proposition **4.** — *Assume that* $1 < \beta < 1 + \min(\omega_s, \omega_u)$. *Then the starting interval* I_0 *is* β*-regular.*

Proof. — Let $\underline{a}^s$ (resp. $\underline{a}^u$) the admissible word in the alphabet $\mathcal{A}$ such that $P_s = P_{\underline{a}^s}$ (resp. $Q_u = Q_{\underline{a}^u}$). By the property (MP2) of Section 2.1, these two words do not have a letter in common.

Let $(P_{\underline{a}}, Q_{\underline{a}}, n)$ an I_0-bicritical rectangle. As $P_{\underline{a}} \cap P_s$ is non-empty, either $\underline{a}$ starts with $\underline{a}^s$ or $\underline{a}^s$ starts with $\underline{a}$. Similarly, either $\underline{a}$ ends with $\underline{a}^u$ or $\underline{a}^u$ ends with $\underline{a}$. As $\underline{a}^s$, $\underline{a}^u$ do not have a common letter, $\underline{a}$ starts with $\underline{a}^s$, ends with $\underline{a}^u$ and $n > n_s + n_u$. In other words, we have

$$(P_{\underline{a}}, Q_{\underline{a}}, n) = (P_s, Q_s, n_s) * (P', Q', n') * (P_u, Q_u, n_u), \tag{5.18}$$

for some $(P', Q', n') \in \mathcal{R}(I_0)$. We conclude that $|P_{\underline{a}}| < |I_0|^{\beta}$, $|Q_{\underline{a}}| < |I_0|^{\beta}$ as required. $\square$

Remark **8.** — The final choice of β is made in Section 9.3. Beside the requirement $1 < \beta < 1 + \min(\omega_s, \omega_u)$ we will ask when $d_s^0 \geqslant d_u^0$ that

$$\beta < \frac{(1 - d_u^0)(d_s^0 + d_u^0)}{d_s^0(2d_s^0 + d_u^0 - 1)} \tag{5.19}$$

(if $d_s^0 \leqslant d_u^0$, exchange d_s^0 and d_u^0). Condition (H4) in Section 1.2 is actually obtained by asking that the right-hand term in the last inequality is larger than 1.

5.6.5.

Proposition **5.** — *Let* $\widetilde{\mathrm{I}} \supset \mathrm{I}$ *be parameter intervals, and let* $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}(\widetilde{\mathrm{I}})$*. If* Q *is* $\widetilde{\mathrm{I}}$*-transverse, then it is also* I*-transverse.*

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

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

Proposition **6.** — *Let* I *be a parameter interval, and let* $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}(\mathrm{I})$*. If* Q *is* I*-transverse, then* P *is* I*-decomposable.*

Proof. — Let us first assume that $\mathrm{Q} \cap \mathrm{Q}_u = \emptyset$. Let $a \in \mathcal{A}$ be such that $\mathrm{Q} \subset \mathrm{R}_a$. We have

$$\mathrm{R}_a \cap \mathrm{W}^s(\Lambda, \widehat{\mathrm{R}}) \subset \bigcup_{(a,a') \in \mathcal{B}} \left(\mathrm{P}_{a,a'} \cap \mathrm{W}^s(\Lambda, \widehat{\mathrm{R}})\right) \cup \mathrm{L}_u; \tag{5.20}$$

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

$$\left(\mathrm{P}(a'), \mathrm{Q}(a'), n+1\right) = (\mathrm{P}, \mathrm{Q}, n) * (\mathrm{P}_{aa'}, \mathrm{Q}_{aa'}, 1), \tag{5.21}$$

and together they form by (5.20) an I-decomposition of P (the canonical one).

Let us now assume that $\mathrm{Q} \subset \mathrm{Q}_u$. As Q is I-transverse, there is an I-decomposition $(\mathrm{P}_\alpha, \mathrm{Q}_\alpha, n_\alpha)_\alpha$ 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 $(\mathrm{P}_\alpha^\pm, \mathrm{Q}_\alpha^\pm, n_\alpha + n + \mathrm{N}_0)$ be the two elements produced by the allowed parabolic composition. Together with the simple children defined by (5.21), they form an I-decomposition of P. □

Corollary **3.** — *Let* I *be a* β*-regular parameter interval and let* $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}(\mathrm{I})$*. If* P *is* I*-critical and* $|\mathrm{P}| > |\mathrm{I}|^\beta$ *or* $|\mathrm{Q}| > |\mathrm{I}|^\beta$ *for some* $t \in \mathrm{I}$*, then* P *is* I*-decomposable.*

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

The decomposability of "fat" critical rectangles is crucial to our analysis.

6. Some properties of the classes $\mathcal{R}(I)$

6.1. *Preliminaries.* — In this section,we start the proof of the existence of the class $\mathcal{R}(I)$. For $I = I_0$, this was done essentially in Section 5.5. Therefore we will, unless specified otherwise, assume that the level of the interval I is > 0. Moreover, the induction hypothesis will guarantee that the classes $\mathcal{R}(\widetilde{I})$, with $\widetilde{I}$ strictly larger than I, have already been constructed with the required properties (R1)–(R7).

Once I is fixed, the construction will be by induction on the length n of an element (P, Q, n). We will denote by $\mathcal{R}_N(I)$ the set of elements (P, Q, n) in $\mathcal{R}(I)$ with $n \leqslant N$. For small N, parabolic composition does not come into play, and therefore the elements of $\mathcal{R}_N(I)$ are those of $\mathcal{R}_N(I_0)$.

Assuming that $\mathcal{R}_{N-1}(I)$ has already been constructed with the required properties, we will consider elements of length N that can be obtained from shorter elements by simple or (allowed) parabolic composition. For these new elements, we will not only prove (R2), (R4), (R7) (in Section 7) as required, but also many other properties that are detailed in the next subsections. The structure theorem of Section 6.5 is particularly important. All these properties are assumed for $\mathcal{R}_{N-1}(I)$, and then they will all be proved for $\mathcal{R}_N(I)$.

6.2. *Children are born from their parent.* — Let $(P, Q, n) \in \mathcal{R}_N(I)$, and let $(\widetilde{P}, \widetilde{Q}, \widetilde{n}) \in \mathcal{R}_{N-1}(I)$ be such that $\widetilde{P}$ is the parent of P. Recall that P is a simple child if $n = \widetilde{n} + 1$, non-simple otherwise. When P is simple, there exists $(a, a') \in \mathcal{B}$ such that (P, Q, n) is the simple composition of $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$ and $(P_{a,a'}, Q_{a,a'}, 1)$.

Proposition **7.** — *Assume that* P *is a non-simple child of* $\widetilde{P}$. *Then there exists* $(P_1, Q_1, n_1) \in \mathcal{R}_{N-1}(I)$ *such that* $\widetilde{Q} \pitchfork_I P_1$ *holds and*

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

Proof. — 1. We first prove that (P, Q, n) can be written as a *parabolic* composition of shorter elements. Otherwise, by (R6), we have

$$(P, Q, n) = (P_0, Q_0, n_0) * (P_1, Q_1, n_1), \tag{6.1}$$

with $n_0, n_1 > 0$. As P is a non-simple child, we must have $n_1 > 1$. Let then $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1)$ be the element of $\mathcal{R}(I)$ such that $\widetilde{P}_1$ is the parent of P_1. If P_1 was a simple child, we would have, for some $(a, a') \in \mathcal{B}$

$$(P_1, Q_1, n_1) = (\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1) * (P_{aa'}, Q_{aa'}, 1), \tag{6.2}$$

$$(P, Q, n) = ((P_0, Q_0, n_0) * (\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1)) * (P_{aa'}, Q_{aa'}, 1), \tag{6.3}$$

in contradiction with the hypothesis that P is a simple child. Therefore, P_1 is a non-simple child; by induction on the length, we can write

$$(P_1, Q_1, n_1) \in (\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1) \boxdot (P_2, Q_2, n_2) \tag{6.4}$$

for some $(P_2, Q_2, n_2) \in \mathcal{R}_{N-2}(I)$ with $\widetilde{Q}_1 \pitchfork_I P_2$.

Let $(\widetilde{P}_0, \widetilde{Q}_0, \widetilde{n}_0) := (P_0, Q_0, n_0) * (\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1)$. We then have $\widetilde{Q}_0 \subset \widetilde{Q}_1$ and hence $\widetilde{Q}_0 \pitchfork_I P_2$ from Proposition 2 in Section 5.6.

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

$$(P, Q, n) \in (\widetilde{P}_0, \widetilde{Q}_0, \widetilde{n}_0) \boxdot (P_2, Q_2, n_2) \tag{6.5}$$

which proves our claim.

2. We next write

$$(P, Q, n) \in (P_0, Q_0, n_0) \boxdot (P_1, Q_1, n_1) \tag{6.6}$$

with n_0 maximal, and want to show that $P_0 = \widetilde{P}$.

Assume that this does not hold, which means $n_0 < \widetilde{n}$. Let (P'_0, Q'_0, n'_0) be the element of $\mathcal{R}_{N-1}(I)$ such that P'_0 is the child of P_0 containing P. As $n_0 < \widetilde{n}$, we have $n'_0 < n$. As $g_t^{n_0}(P) \subset L_u$ by (6.6), P'_0 must be a non-simple child. Then, from the induction hypothesis, we can write

$$(P'_0, Q'_0, n'_0) \in (P_0, Q_0, n_0) \boxdot (P'_1, Q'_1, n'_1) \tag{6.7}$$

for some $(P'_1, Q'_1, n'_1) \in \mathcal{R}(I)$ with $Q_0 \pitchfork_I P'_1$, $P'_1 \supset P_1$. Let $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1)$ be the element of $\mathcal{R}_{N-2}(I)$ such that $\widetilde{P}_1$ is the parent of P_1 We have thus $P'_1 \supset \widetilde{P}_1$. As Q_0 is I-transverse to P'_1, it is also I-transverse to $\widetilde{P}_1$ (Proposition 2), and we can define $(\widetilde{P}_0, \widetilde{Q}_0, \widetilde{n}_0) \in \mathcal{R}_{N-1}(I)$ by

$$(\widetilde{P}_0, \widetilde{Q}_0, \widetilde{n}_0) \in (P_0, Q_0, n_0) \boxdot (\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1), \tag{6.8}$$

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

$$(P_1, Q_1, n_1) \in (\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1) \boxdot (P_2, Q_2, n_2) \tag{6.9}$$

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

$$(P, Q, n) \in (\widetilde{P}_0, \widetilde{Q}_0, \widetilde{n}_0) \boxdot (P_2, Q_2, n_2) \tag{6.10}$$

which contradicts the maximality of n_0. The proof is complete. □

6.3. *Coherence and concavity.*

6.3.1. We present together the following two properties of $\mathcal{R}(I)$ because the proofs are interwoven.

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

Proposition **8** *(Coherence). — Let* $(P, Q, n) \in \mathcal{R}_N(I)$*,* $\widetilde{I}$ *be the parent of* I*, and* $(P', Q', n') \in \mathcal{R}(\widetilde{I})$*. If* $P' \subset P$*, then* (P, Q, n) *belongs to* $\mathcal{R}(\widetilde{I})$*.*

The coherence property implies that the parent-child relationship is independent of the parameter interval under consideration, provided the child is already constructed.

Proposition **9** *(Concavity). — Let* $\widetilde{I}$ *be the parent of* I*.*

1. *Let* (P'_0, Q'_0, n'_0)*,* (P'_1, Q'_1, n'_1) *be elements of* $\mathcal{R}_N(I)$*,* (P_0, Q_0, n_0)*,* (P_1, Q_1, n_1) *be elements of* $\mathcal{R}_N(I) \cup \mathcal{R}(\widetilde{I})$ *(with possibly* n_0 *or* $n_1 > N$*) such that*
$$Q_0 \subset Q'_0 \subset Q_u, \qquad P_1 \subset P'_1 \subset P_s.$$
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.*
2. *Let* I' *be a parameter interval containing* $\widetilde{I}$*, let* (P'_0, Q'_0, n'_0)*,* (P_1, Q_1, n_1) *be elements of* $\mathcal{R}_N(I')$*,* (P_0, Q_0, n_0) *an element of* $\mathcal{R}(I')$ *(with possibly* $n_0 > N$*) with*
$$Q_0 \subset Q'_0 \subset Q_u, \qquad P_1 \subset P_s.$$
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.*
3. *Let* I' *be a parameter interval containing* $\widetilde{I}$*, let* (P_0, Q_0, n_0)*,* (P'_1, Q'_1, n'_1) *be elements of* $\mathcal{R}_N(I')$*,* (P_1, Q_1, n_1) *an element of* $\mathcal{R}(I')$ *(with possibly* $n_1 > N$*) with*
$$P_1 \subset P'_1 \subset P_s, \qquad Q_0 \subset Q_u.$$
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.*

The concavity property is very helpful in the sequel. The proof of the proposition will help to explain why the definition of the transversality relation had to be complicated.

In the following subsections, we will successively:

– prove Proposition 8 for $\mathcal{R}_N(I)$, assuming Proposition 8 and Proposition 9 for $\mathcal{R}_{N-1}(I)$;
– prove Proposition 9 for $\mathcal{R}_N(I)$, assuming Proposition 8 for $\mathcal{R}_N(I)$.

6.3.2. *Proof of Proposition 8.* *Proof.* — Let $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$ be the element of $\mathcal{R}(\widetilde{I})$ such that $\widetilde{P}$ is the parent of P′ in $\mathcal{R}(\widetilde{I})$. By enlarging P′ if necessary, one can assume that $P' \subset P \subset \widetilde{P}$, $P \neq \widetilde{P}$.

If P′ is a simple child, we have $P = P'$ and we are done. Assume that P′ is a non-simple child of $\widetilde{P}$.

Let $(\widehat{P}, \widehat{Q}, \widehat{n})$ be the element of $\mathcal{R}_N(I)$ such that $\widehat{P}$ is the child of $\widetilde{P}$ in $\mathcal{R}(I)$ containing P; as $\widehat{P}$ also contains P′ it has to be a non-simple child.

Applying Proposition 7 twice, we find $(P_1, Q_1, n_1) \in \mathcal{R}_{N-1}(I)$, $(P'_1, Q'_1, n'_1) \in \mathcal{R}(\widetilde{I})$ such that $\widetilde{Q} \pitchfork_I P_1$, $\widetilde{Q} \pitchfork_{\widetilde{I}} P'_1$ both hold and

$$(\widehat{P}, \widehat{Q}, \widehat{n}) \in (\widetilde{P}, \widetilde{Q}, \widetilde{n}) \,\square\, (P_1, Q_1, n_1), \tag{6.11}$$

$$(P', Q', n') \in (\widetilde{P}, \widetilde{Q}, \widetilde{n}) \,\square\, (P'_1, Q'_1, n'_1). \tag{6.12}$$

As $P' \subset P \subset \widehat{P}$, we have $P'_1 \subset P_1$. By Proposition 8 for $\mathcal{R}_{N-1}(I)$, we have $(P_1, Q_1, n_1) \in \mathcal{R}(\widetilde{I})$. By Proposition 9, part 3, for $\mathcal{R}_{N-1}(I)$, we have $\widetilde{Q} \pitchfork_{\widetilde{I}} P_1$ and $(\widehat{P}, \widehat{Q}, \widehat{n}) \in \mathcal{R}(\widetilde{I})$.

As $P' \subset P \subset \widehat{P}$ and P′ is a child of $\widetilde{P}$ in $\mathcal{R}(\widetilde{I})$, we must have $P' = P = \widehat{P}$ and the proof is complete. □

6.3.3. *Proof of Proposition 9.*

1. With (P_0, Q_0, n_0), (P'_0, Q'_0, n'_0), (P_1, Q_1, n_1), (P'_1, Q'_1, n'_1) as in the first part of Proposition 9, we assume that both $Q_0 \pitchfork_I P'_1$ and $Q'_0 \pitchfork_I P_1$ hold.

   By definition of the transversality relation, there exist parameter intervals I_1, I_2 containing I, elements $(\widetilde{P}_0, \widetilde{Q}_0, \widetilde{n}_0) \in \mathcal{R}(I_1)$, $(\widetilde{P}'_1, \widetilde{Q}'_1, \widetilde{n}'_1) \in \mathcal{R}_N(I_1)$, $(\widetilde{P}'_0, \widetilde{Q}'_0, \widetilde{n}'_0) \in \mathcal{R}_N(I_2)$, $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1) \in \mathcal{R}(I_2)$ such that $Q_0 \subset \widetilde{Q}_0$, $P'_1 \subset \widetilde{P}'_1$, $Q'_0 \subset \widetilde{Q}'_0$, $P_1 \subset \widetilde{P}_1$ and

$$\widetilde{Q}_0 \,\overline{\pitchfork}_{I_1}\, \widetilde{P}'_1, \tag{6.13}$$

$$\widetilde{Q}'_0 \,\overline{\pitchfork}_{I_2}\, \widetilde{P}_1. \tag{6.14}$$

   If we have either $Q'_0 \subset \widetilde{Q}_0$ or $P'_1 \subset \widetilde{P}_1$, we can already conclude that $Q'_0 \pitchfork_I P'_1$. Assume therefore that $\widetilde{Q}_0 \subset Q'_0$, $\widetilde{P}_1 \subset P'_1$. Assume also for instance that $I_1 \subset I_2$. We have $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1) \in \mathcal{R}(I_2)$ and $(\widetilde{P}'_1, \widetilde{Q}'_1, \widetilde{n}'_1) \in \mathcal{R}_N(I_1)$. By the coherence property (Proposition 8) for $\mathcal{R}_N(I_1)$, the element $(\widetilde{P}'_1, \widetilde{Q}'_1, \widetilde{n}'_1)$ belongs to $\mathcal{R}(I_2)$. We will show that

$$\widetilde{Q}'_0 \,\overline{\pitchfork}_{I_2}\, \widetilde{P}'_1 \tag{6.15}$$

   which implies that Q'_0 and P'_1 are I-transverse.

   We check properties (T1)–(T3) of Section 5.4. For all $t \in I_2$, by the estimate (3.49) of Section 3.6 and (T1) for $\widetilde{Q}'_0 \,\overline{\pitchfork}_{I_2}\, \widetilde{P}_1$, we have

$$\delta_{LR}(\widetilde{Q}'_0, \widetilde{P}'_1) \geqslant \delta_{LR}(\widetilde{Q}'_0, \widetilde{P}_1) \geqslant 2|I_2|. \tag{6.16}$$

Next, by (3.47) and (T2) for $\widetilde{Q}'_0 \,\overline{\pitchfork}_{I_2}\, \widetilde{P}_1$, there exists $t_0 \in I_2$ such that

$$\delta_R(\widetilde{Q}'_0, \widetilde{P}'_1) \geqslant \delta_R(\widetilde{Q}'_0, \widetilde{P}_1) \geqslant 2|\widetilde{Q}'_0|^{1-\eta}. \tag{6.17}$$

Finally, by (3.48) and (T3) for $\widetilde{Q}_0 \,\overline{\pitchfork}_{I_1}\, \widetilde{P}'_1$, there exists $t_1 \in I_1 \subset I_2$ such that

$$\delta_L(\widetilde{Q}'_0, \widetilde{P}'_1) \geqslant \delta_L(\widetilde{Q}_0, \widetilde{P}'_1) \geqslant 2|\widetilde{P}'_1|^{1-\eta}. \tag{6.18}$$

We have proved (6.15) and this concludes the proof of the first statement in the proposition.

2. We assume now that, with I', (P_0, Q_0, n_0), (P'_0, Q'_0, n'_0), (P_1, Q_1, n_1) as in the second part of the proposition, both $Q_0 \pitchfork_{I'} P_1$ and $Q'_0 \pitchfork_I P_1$ hold.

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

$$\widetilde{Q}_0 \,\overline{\pitchfork}_{\widehat{I'}}\, \widetilde{P}_1, \tag{6.19}$$

$$\widetilde{Q}'_0 \,\overline{\pitchfork}_{\widehat{I}}\, \widetilde{P}'_1 \tag{6.20}$$

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

   We claim that (P^*_1, Q^*_1, n^*_1) always belongs to $\mathcal{R}(\widehat{I'})$. This is clear if $\widetilde{P}'_1 \subset \widetilde{P}_1$; on the other hand, if $\widetilde{P}_1 \subset \widetilde{P}'_1$, it follows from coherence (Proposition 8) for $\mathcal{R}_N(\widehat{I})$. As $\widetilde{Q}_0 \subset \widetilde{Q}'_0$, it also follows from coherence for $\mathcal{R}_N(\widehat{I})$ that $(\widetilde{P}'_0, \widetilde{Q}'_0, \widetilde{n}'_0)$ belongs to $\mathcal{R}(\widehat{I'})$.

   We will show that

$$\widetilde{Q}'_0 \,\overline{\pitchfork}_{\widehat{I'}}\, P^*_1 \tag{6.21}$$

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

   We check properties (T1)–(T3) of Section 5.4. For all $t \in \widehat{I'}$, we have, from (3.49) and (T1) for $\widetilde{Q}_0 \,\overline{\pitchfork}_{\widehat{I'}}\, \widetilde{P}_1$

$$\delta_{LR}(\widetilde{Q}'_0, P^*_1) \geqslant \delta_{LR}(\widetilde{Q}_0, \widetilde{P}_1) \geqslant 2|\widehat{I'}|. \tag{6.22}$$

   From (T2) for $\widetilde{Q}'_0 \,\overline{\pitchfork}_{\widehat{I}}\, \widetilde{P}'_1$, there exists $t_0 \in \widehat{I} \subset I' \subset \widehat{I'}$ such that

$$\delta_R(\widetilde{Q}'_0, \widetilde{P}'_1) \geqslant 2|\widetilde{Q}'_0|^{1-\eta}. \tag{6.23}$$

Then, by (3.47), for the same t_0, we have

$$\delta_R(\widetilde{Q}'_0, P^*_1) \geqslant \delta_R(\widetilde{Q}'_0, \widetilde{P}'_1) \geqslant 2|\widetilde{Q}'_0|^{1-\eta}. \tag{6.24}$$

When $P_1^* = \widetilde{P}'_1$, it follows directly from (T3) for $\widetilde{Q}'_0 \,\overline{\pitchfork}_{\widehat{I}}\, \widetilde{P}'_1$ that we have

$$\delta_L(\widetilde{Q}'_0, P_1^*) \geqslant 2|P_1^*|^{1-\eta} \tag{6.25}$$

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

When $P_1^* = \widetilde{P}_1$, we use (T3) for $\widetilde{Q}_0 \,\overline{\pitchfork}_{\widehat{I}'}\, \widetilde{P}_1$ and (3.48) to find $t_1 \in \widehat{I}'$ such that

$$\delta_L(\widetilde{Q}'_0, P_1^*) \geqslant \delta_L(\widetilde{Q}_0, \widetilde{P}_1) \geqslant 2|P_1^*|^{1-\eta}. \tag{6.26}$$

We have thus proved (6.21). This proves the second part of the proposition. The third part is proven in a symmetric way, exchanging P's and Q's.

6.3.4. *Further forms of concavity.* — One obtains more general statements than in Proposition 9 by combining its different parts.

Corollary **4.** — *Let* I' *be a parameter interval containing the parent of* I, *and let* (P'_0, Q'_0, n'_0), (P'_1, Q'_1, n'_1), (P_0, Q_0, n_0), (P_1, Q_1, n_1) *be such that*

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

1. *Assume that* $(P'_0, Q'_0, n'_0), (P'_1, Q'_1, n'_1) \in \mathcal{R}_N(I)$, *and that* $(P_0, Q_0, n_0), (P_1, Q_1, n_1) \in \mathcal{R}(I')$. *If both* $Q_0 \pitchfork_{I'} P_1$ *and* $Q'_0 \pitchfork_I P'_1$ *hold, and* $\min(n_0, n_1) \leqslant N$, *then* $Q'_0 \pitchfork_{I'} P'_1$ *also holds.*
2. *Assume that* (P_0, Q_0, n_0), $(P'_1, Q'_1, n'_1) \in \mathcal{R}_N(I)$, *and that* $(P'_0, Q'_0, n'_0) \in \mathcal{R}_N(I')$, $(P_1, Q_1, n_1) \in \mathcal{R}(I')$ *(with possibly* $n_1 > N$*). 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.*
3. *Assume that* (P'_0, Q'_0, n'_0), $(P_1, Q_1, n_1) \in \mathcal{R}_N(I)$, *and that* $(P'_1, Q'_1, n'_1) \in \mathcal{R}_N(I')$, $(P_0, Q_0, n_0) \in \mathcal{R}(I')$ *(with possibly* $n_0 > N$*). 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. — 1. For the first statement of the corollary, we first observe that, by coherence for $\mathcal{R}_N(I)$, the elements (P'_0, Q'_0, n'_0), (P'_1, Q'_1, n'_1) belong to $\mathcal{R}(I')$.

Assume for instance that $n_0 \leqslant N$. From $Q'_0 \pitchfork_I P'_1$ and Proposition 2, we have $Q_0 \pitchfork_I P'_1$. Then, from $Q_0 \pitchfork_I P'_1$ and $Q_0 \pitchfork_{I'} P_1$, we have $Q_0 \pitchfork_{I'} P'_1$ by Proposition 9 for $\mathcal{R}_N(I)$. Finally, from $Q_0 \pitchfork_{I'} P'_1$ and $Q'_0 \pitchfork_I P'_1$, we have $Q'_0 \pitchfork_{I'} P'_1$ again by Proposition 9 for $\mathcal{R}_N(I)$.

2. For the second statement of the corollary, we first observe that, by coherence for $\mathcal{R}_N(I)$, the element (P'_1, Q'_1, n'_1) belong to $\mathcal{R}(I')$.

From $Q'_0 \pitchfork_{I'} P_1$ and Proposition 2, we have $Q_0 \pitchfork_{I'} P'_1$. Then, from $Q_0 \pitchfork_I P'_1$ and $Q_0 \pitchfork_{I'} P_1$, we have $Q_0 \pitchfork_{I'} P'_1$ by Proposition 9 for $\mathcal{R}_N(I)$. Finally, from $Q_0 \pitchfork_{I'} P'_1$ and $Q'_0 \pitchfork_{I'} P_1$, we have $Q'_0 \pitchfork_{I'} P'_1$ again by Proposition 9 for $\mathcal{R}_N(I)$.

3. The proof of the third statement of the corollary is the same than for the second, exchanging P's and Q's. □

6.4. *Another transversality criterion.*

Proposition **10.** — *Let* (P_0, Q_0, n_0), (P_1, Q_1, n_1), (P'_1, Q'_1, n'_1) *be elements of* $\mathcal{R}_N(I)$ *such that* $Q_0 \subset Q_u$ *and* $P_1 \subset P'_1 \subset P_s$. *Assume that* $Q_0 \pitchfork_I P_1$ *holds and that* $2|P'_1|^{1-\eta} \leqslant |I|$ *for some* $t_1 \in I$. *Then* Q_0 *and* P'_1 *are also* I-*transverse.*

Proof. — By definition of the transversality relation, there exist $\widetilde{I} \supset I$, $(\widetilde{P}_0, \widetilde{Q}_0, \widetilde{n}_0)$, $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1) \in \mathcal{R}(\widetilde{I})$ such that $Q_0 \subset \widetilde{Q}_0$, $P_1 \subset \widetilde{P}_1$ and $\widetilde{Q}_0 \,\overline{\pitchfork}_{\widetilde{I}}\, \widetilde{P}_1$.

If $P'_1 \subset \widetilde{P}_1$ this already implies that $Q_0 \pitchfork_I P'_1$. Let us assume that $\widetilde{P}_1 \subset P'_1$. We will show that $\widetilde{Q}_0 \,\overline{\pitchfork}_{\widetilde{I}}\, P'_1$ holds. By coherence (Proposition 8), we have $(P'_1, Q'_1, n'_1) \in \mathcal{R}(\widetilde{I})$.

Let us check (T1)–(T3).

By (T1) for $Q_0 \pitchfork_I P_1$ and (3.49) in Section 3.6, we have, for all $t \in \widetilde{I}$:

$$\delta_{LR}(\widetilde{Q}_0, P'_1) \geqslant \delta_{LR}(\widetilde{Q}_0, \widetilde{P}_1) \geqslant 2|\widetilde{I}|. \tag{6.27}$$

By (T2) for $Q_0 \pitchfork_I P_1$ and (3.47), there exists $t_0 \in \widetilde{I}$ such that

$$\delta_R(\widetilde{Q}_0, P'_1) \geqslant \delta_R(\widetilde{Q}_0, \widetilde{P}_1) \geqslant 2|\widetilde{Q}_0|^{1-\eta}. \tag{6.28}$$

Finally, we have, for all $t \in I$, by (3.39)

$$\begin{aligned} \delta_L(\widetilde{Q}_0, P'_1) &\geqslant \delta_{LR}(\widetilde{Q}_0, P'_1) - C|P'_1| \\ &\geqslant 2|I| - C|P'_1|. \end{aligned} \tag{6.29}$$

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

$$2|I| - C|P'_1| \geqslant 4|P'_1|^{1-\eta} - C|P'_1| \geqslant 2|P'_1|^{1-\eta}. \tag{6.30}$$

We have shown that $\widetilde{Q}_0 \,\overline{\pitchfork}_{\widetilde{I}}\, P'_1$ holds, and this implies that $Q_0 \pitchfork_I P'_1$ holds as required. □

6.5. *A structure theorem for new rectangles.*

6.5.1. *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 Proposition 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

$$(P, Q, n) \in (P_0, Q_0, n_0) \square (P_1, Q_1, n_1) \square (P_2, Q_2, n_2). \tag{6.31}$$

The same considerations extend immediately, by induction on k, to the case of elements $(P_0, Q_0, n_0), \dots, (P_k, Q_k, n_k)$ such that $P_i \subset P_s$ for $0 < i \leq k$, $Q_i \subset Q_u$ 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):

$$(P, Q, n) \in (P_0, Q_0, n_0) \square \cdots \square (P_k, Q_k, n_k). \tag{6.32}$$

6.5.2. *Statement of the structure theorem.* — We have seen in Section 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 **2.** — *Let* I *be a parameter interval of level* > 0, $\widetilde{I}$ *be the parent interval, and let* (P, Q, n) *be an element of* $\mathcal{R}_N(I)$ *which is not (the restriction of) an element of* $\mathcal{R}(\widetilde{I})$*. Then there exists* $k > 0$*, elements* $(P_0, Q_0, n_0), \dots, (P_k, Q_k, n_k)$ *of* $\mathcal{R}(\widetilde{I})$ *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* $0 \leq i < k$, $Q_i \pitchfork_{\widetilde{I}} P_{i+1}$ *does not hold for* $0 \leq i < k$ *and*

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

Moreover, these elements are uniquely determined by these conditions, P_i *is* $\widetilde{I}$*-critical for* $0 < i \leq k$ *and* Q_i *is* $\widetilde{I}$*-critical for* $0 \leq i < k$*. The rectangle* P_0 *is the thinnest* $\widetilde{I}$*-defined vertical rectangle containing* P*, the rectangle* Q_k *is the thinnest* $\widetilde{I}$*-defined horizontal rectangle containing* Q*.*

The rest of Section 6.5 is devoted to the proof of the theorem.

6.5.3. *Proof.* — 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 $\widetilde{I}$-interval if there exists $(\widetilde{P}, \widetilde{Q}, \widetilde{n}) \in \mathcal{R}(\widetilde{I})$ such that

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

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

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

(a) $p = m'$.

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

$$(\widetilde{\mathrm{P}}'', \widetilde{\mathrm{Q}}'', \widetilde{n}'') := (\widetilde{\mathrm{P}}, \widetilde{\mathrm{Q}}, \widetilde{n}) * (\widetilde{\mathrm{P}}', \widetilde{\mathrm{Q}}', \widetilde{n}') \tag{6.33}$$

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

(b) $p > m'$.

Then, $\widetilde{\mathrm{P}}$ is not a simple child of $\widehat{\mathrm{P}}$, because otherwise we would have $\widehat{n} = \widetilde{n} - 1 \geq m' - m$. By Proposition 7, there exists $(\widetilde{\mathrm{P}}_0, \widetilde{\mathrm{Q}}_0, \widetilde{n}_0)$ in $\mathcal{R}(\widetilde{\mathrm{I}})$ such that

$$(\widetilde{\mathrm{P}}, \widetilde{\mathrm{Q}}, \widetilde{n}) \in (\widehat{\mathrm{P}}, \widehat{\mathrm{Q}}, \widehat{n}) \,\square\, (\widetilde{\mathrm{P}}_0, \widetilde{\mathrm{Q}}_0, \widetilde{n}_0). \tag{6.34}$$

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

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 $\widetilde{\mathrm{I}}$-intervals $[\widehat{m}, p]$ and $[m', p']$ with $m < \widehat{m} \leq m'$ and hence by induction $[\widehat{m}, p']$ is an $\widetilde{\mathrm{I}}$-interval. Let $(\widetilde{\mathrm{P}}_1, \widetilde{\mathrm{Q}}_1, \widetilde{n}_1)$ be the corresponding element of $\mathcal{R}(\widetilde{\mathrm{I}})$; we have $g_t^{\widehat{m}}(\mathrm{P}) \subset \widetilde{\mathrm{P}}_1 \subset \widetilde{\mathrm{P}}_0$. From (6.34), $\widehat{\mathrm{Q}} \pitchfork_{\widetilde{\mathrm{I}}} \widetilde{\mathrm{P}}_0$ holds, hence also does $\widehat{\mathrm{Q}} \pitchfork_{\widetilde{\mathrm{I}}} \widetilde{\mathrm{P}}_1$ by Proposition 2. Then, the parabolic composition of $(\widehat{\mathrm{P}}, \widehat{\mathrm{Q}}, \widehat{n})$ and $(\widetilde{\mathrm{P}}_1, \widetilde{\mathrm{Q}}_1, \widetilde{n}_1)$ is allowed and defines an element of $\mathcal{R}(\widetilde{\mathrm{I}})$ which guarantees that $[m, p']$ is an $\widetilde{\mathrm{I}}$-interval. □

6.5.4. We will now show that the $(\mathrm{P}_i, \mathrm{Q}_i, n_i)$ in the theorem are uniquely determined by their properties. Indeed, define $m_0 = 0$, $p_0 = n_0$ and for $i > 0$:

$$m_i = p_{i-1} + \mathrm{N}_0, \qquad p_i = m_i + n_i. \tag{6.35}$$

Lemma **2.** — *The maximal* $\widetilde{\mathrm{I}}$*-intervals are exactly the* $[m_i, p_i]$, $0 \leq i \leq k$, *with associated elements* $(\mathrm{P}_i, \mathrm{Q}_i, n_i)$.

Proof. — First, the $[m_i, p_i]$ are indeed $\widetilde{\mathrm{I}}$-intervals with associated elements $(\mathrm{P}_i, \mathrm{Q}_i, n_i)$. To complete the proof, it is sufficient to show that no $\widetilde{\mathrm{I}}$-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 $(\widetilde{\mathrm{P}}, \widetilde{\mathrm{Q}}, \widetilde{n})$ and minimal $\widetilde{n} = p - m$. As $g_t^{\ell}(\mathrm{P}) \subset g_t^{\ell - p_i}(\mathrm{L}_u)$ does not intersect R for $p_i < \ell < m_{i+1}$, we must have $m \leqslant p_i$ and $m_{i+1} \leqslant p$. By property (R6) of $\mathcal{R}(\widetilde{\mathrm{I}})$ (Section 5.3) and the minimality of $\widetilde{n}$, there exists $(\widetilde{\mathrm{P}}_0, \widetilde{\mathrm{Q}}_0, \widetilde{n}_0)$, $(\widetilde{\mathrm{P}}_1, \widetilde{\mathrm{Q}}_1, \widetilde{n}_1)$ in $\mathcal{R}(\widetilde{\mathrm{I}})$ such that

$$(\widetilde{\mathrm{P}}, \widetilde{\mathrm{Q}}, \widetilde{n}) \in (\widetilde{\mathrm{P}}_0, \widetilde{\mathrm{Q}}_0, \widetilde{n}_0) \,\square\, (\widetilde{\mathrm{P}}_1, \widetilde{\mathrm{Q}}_1, \widetilde{n}_1) \tag{6.36}$$

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

6.5.5. Lemma 2 allows us to *define* k as being the number of maximal $\widetilde{I}$-intervals minus one, and to define the $(P_i, Q_i, n_i) \in \mathcal{R}(\widetilde{I})$ as the elements of $\mathcal{R}(\widetilde{I})$ associated to the successive maximal $\widetilde{I}$-intervals. Observe that the maximal $\widetilde{I}$-intervals $[m_i, p_i]$, $(0 \leqslant i \leqslant k)$ must indeed satisfy $m_0 = 0$, $m_{i+1} = p_i + N_0$ for $0 \leqslant i < k$: every $\ell \in [0, n]$ not contained in an $\widetilde{I}$-interval is such that $g_t^{\ell - N_1}(P) \subset L_u$ for some $0 < N_1 < N_0$ and then no $\widetilde{I}$-interval intersects with $(\ell - N_1, \ell - N_1 + N_0)$, while the degenerates intervals$\{\ell - N_1\}$, $\{\ell - N_1 + N_0\}$ are $\widetilde{I}$-intervals. We observe also that $Q_i \pitchfork_{\widetilde{I}} P_{i+1}$ does not hold because otherwise $[m_i, p_{i+1}]$ would be an $\widetilde{I}$-interval.

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

$$(P^{(i)}, Q^{(i)}, p_i) \in (P_0, Q_0, n_0) \boxdot \cdots \boxdot (P_i, Q_i, n_i) \tag{6.37}$$

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).

6.5.7. We first prove that Q_i is $\widetilde{I}$-critical. Assume by contradiction that Q_i is $\widetilde{I}$-transverse. Then P_i is $\widetilde{I}$-decomposable. Let then $(\widehat{P}_i, \widehat{Q}_i, \widehat{n}_i)$ be an element of $\mathcal{R}(\widetilde{I})$ such that $\widehat{P}_i$ is a child of P_i intersecting $g_t^{m_i}(P \cap \Lambda)$. Let $(\widehat{P}^{(i)}, \widehat{Q}^{(i)}, \widehat{n}^{(i)})$ be the element of $\mathcal{R}_N(I)$ such that $\widehat{P}^{(i)}$ is the child of $P^{(i)}$ containing P. Both $\widehat{P}_i$ and $\widehat{P}^{(i)}$ are non-simple children by Section 6.5.5.

We apply Proposition 7 twice. We find $(\widetilde{P}_{i+1}, \widetilde{Q}_{i+1}, \widetilde{n}_{i+1})$ in $\mathcal{R}_{N-1}(I)$, $(\widetilde{P}'_{i+1}, \widetilde{Q}'_{i+1}, \widetilde{n}'_{i+1})$ in $\mathcal{R}(\widetilde{I})$ such that both $Q_i \pitchfork_{\widetilde{I}} \widetilde{P}'_{i+1}$ and $Q^{(i)} \pitchfork_I \widetilde{P}_{i+1}$ hold and

$$(\widehat{P}_i, \widehat{Q}_i, \widehat{n}_i) \in (P_i, Q_i, n_i) \boxdot (\widetilde{P}'_{i+1}, \widetilde{Q}'_{i+1}, \widetilde{n}'_{i+1}), \tag{6.38}$$

$$(\widehat{P}^{(i)}, \widehat{Q}^{(i)}, \widehat{n}^{(i)}) \in (P^{(i)}, Q^{(i)}, p_i) \boxdot (\widetilde{P}_{i+1}, \widetilde{Q}_{i+1}, \widetilde{n}_{i+1}). \tag{6.39}$$

We must have $m_i + \widehat{n}_i > n$, because otherwise $[m_i, m_i + \widehat{n}_i]$ would be an $\widetilde{I}$-interval strictly larger than $[m_i, p_i]$. Thus we have $\widetilde{n}'_{i+1} > \widetilde{n}_{i+1}$. But then, from $Q_i \pitchfork_{\widetilde{I}} \widetilde{P}'_{i+1}$ and $Q^{(i)} \pitchfork_I \widetilde{P}_{i+1}$, we deduce by Corollary 4 (part 2) that $Q_i \pitchfork_{\widetilde{I}} \widetilde{P}_{i+1}$. Parabolic composition yields an element $(\overline{P}_i, \overline{Q}_i, \overline{n}_i) \in \mathcal{R}(\widetilde{I})$ with $P_i \supsetneq \overline{P}_i \supsetneq \hat{P}_i$, in contradiction with the definition of $\widehat{P}_i$.

6.5.8. The proof that P_{i+1} is $\widetilde{I}$-critical is rather similar. We assume by contradiction that it is $\widetilde{I}$-transverse. Then Q_{i+1} is $\widetilde{I}$-decomposable and we can find $(P^*_{i+1}, Q^*_{i+1}, n^*_{i+1}) \in \mathcal{R}(\widetilde{I})$ such that Q^*_{i+1} is a child of Q_{i+1} intersecting $g^{p_{i+1}}(P \cap \Lambda)$. It is a non-simple child as $g^{p_{i+1}}(P \cap \Lambda)$ is contained in L_u. By Proposition 7, there exists $(P_*^{(i)}, Q_*^{(i)}, n_*^{(i)}) \in \mathcal{R}(\widetilde{I})$ such that $Q_*^{(i)} \pitchfork_{\widetilde{I}} P_{i+1}$ holds and

$$(\mathbf{6.40}) \qquad (P_{i+1}^*, Q_{i+1}^*, n_{i+1}^*) \in (P_*^{(i)}, Q_*^{(i)}, n_*^{(i)}) \,\square\, (P_{i+1}, Q_{i+1}, n_{i+1}).$$

We must have $n_*^{(i)} > p_i$, because otherwise $[p_i - n_*^{(i)}, p_{i+1}]$ would be an $\widetilde{I}$-interval strictly larger than $[m_{i+1}, p_{i+1}]$. Then, by coherence for $\mathcal{R}_N(I)$, we have $(P^{(i)}, Q^{(i)}, p_i) \in \mathcal{R}(\widetilde{I})$ and thus $i = 0$.

Let $(\widehat{P}^{(i)}, \widehat{Q}^{(i)}, \widehat{n}^{(i)})$, $(\widetilde{P}_{i+1}, \widetilde{Q}_{i+1}, \widetilde{n}_{i+1})$ be as in Section 6.5.7. From $Q_*^{(i)} \pitchfork_{\widetilde{I}} P_{i+1}$ and $Q^{(i)} \pitchfork_I \widetilde{P}_{i+1}$, we deduce by Corollary 4 (part 2) (if $\widetilde{n}_{i+1} \geqslant n_{i+1}$) or Proposition 9 (if $\widetilde{n}_{i+1} \leqslant n_{i+1}$) that $Q^{(i)} \pitchfork_{\widetilde{I}} P_{i+1}$. This means that $Q_0 \pitchfork_{\widetilde{I}} P_1$ and we conclude that $[0, p_1]$ is an $\widetilde{I}$-interval, a contradiction.

6.5.9. We now prove that $Q^{(i)}$ and P_{i+1} are I-transverse. Let $(\widehat{P}^{(i)}, \widehat{Q}^{(i)}, \widehat{n}^{(i)})$, $(\widetilde{P}_{i+1}, \widetilde{Q}_{i+1}, \widetilde{n}_{i+1})$ be as in Section 6.5.7. If $\widetilde{n}_{i+1} \leq n_{i+1}$, we have $Q^{(i)} \pitchfork_I \widetilde{P}_{i+1}$ by (6.39) and thus also $Q^{(i)} \pitchfork_I P_{i+1}$ by Proposition 2. Let us assume that $\widetilde{n}_{i+1} > n_{i+1}$.

We claim that, under this hypothesis, Q_{i+1} is $\widetilde{I}$-critical. Indeed, if it was $\widetilde{I}$-transverse, P_{i+1} would be $\widetilde{I}$-decomposable and we would find an element $(\widehat{P}_{i+1}, \widehat{Q}_{i+1}, \widehat{n}_{i+1})$ of $\mathcal{R}(\widetilde{I})$ such that $\widehat{P}_{i+1}$ is a child of P_{i+1} intersecting $g^{m_{i+1}}(P \cap \Lambda)$. By coherence (Proposition 8), we should have $\widehat{P}_{i+1} \supset \widetilde{P}_{i+1}$ and $[m_{i+1}, m_{i+1} + \widehat{n}_{i+1}]$ would be an $\widetilde{I}$-interval larger than $[m_{i+1}, p_{i+1}]$, a contradiction which proves our claim.

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

$$(\mathbf{6.41}) \qquad |P_{i+1}| < |\widetilde{I}|^\beta$$

and thus also (with ε_0 small enough)

$$(\mathbf{6.42}) \qquad 2|P_{i+1}|^{1-\eta} < |I|.$$

It now follows from Proposition 10 and $Q^{(i)} \pitchfork_I \widetilde{P}_{i+1}$ that $Q^{(i)}$ and P_{i+1} are I-transverse.

6.5.10. We finally show that Q_i and P_{i+1} are I-transverse. When $i = 0$, we have $Q^{(0)} = Q_0$, so this has been done in Section 6.5.9.

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

$$(\mathbf{6.43}) \qquad |Q_i| < |\widetilde{I}|^\beta,$$

and thus also

$$(\mathbf{6.44}) \qquad 2|Q_i|^{1-\eta} < |I|.$$

It follows from Proposition 10 and $Q^{(i)} \pitchfork_I P_{i+1}$ that Q_i and P_{i+1} are I-transverse.

To conclude the induction step of Section 6.5.6, 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 last assertion in the theorem follows from Section 6.5.5. The proof of the theorem is now complete.

6.6. *Width estimates.*

6.6.1. I-*special rectangles.* — As long as $\mathcal{R}(I)$ has not been fully constructed, we cannot decide whether a rectangle P or Q is I-critical or not. On the other hand, we can decide whether $\tilde{P}$ or $\tilde{Q}$ is $\tilde{I}$-critical or not when $\tilde{I}$ strictly contains I and $(\tilde{P}, \tilde{Q}, \tilde{n}) \in \mathcal{R}(\tilde{I})$. The following definition is therefore useful.

Let $(P, Q, n) \in \mathcal{R}_N(I)$. We say that P is I-*special* if either $(P, Q, n) \in \mathcal{R}(I_0)$ (i.e. no parabolic composition is ever involved in the construction of P) or, denoting by $\tilde{I}$ the parent of I and by $(\tilde{P}, \tilde{Q}, \tilde{n})$ the element of $\mathcal{R}(\tilde{I})$ such that $P \subset \tilde{P}$ and $\tilde{P}$ is smallest possible, we have that $\tilde{P}$ is $\tilde{I}$-critical. We define similarly the property for Q.

There are therefore three cases:

- $(P, Q, n) \in \mathcal{R}(I_0)$; in this case, P is always I-special.
- $(P, Q, n) \in \mathcal{R}(\tilde{I}) - \mathcal{R}(I_0)$; in this case, P is I-special iff it is $\tilde{I}$-critical.
- $(P, Q, n) \in \mathcal{R}(I) - \mathcal{R}(\tilde{I})$; in this case, we apply the structure theorem in the last subsection and see that $\tilde{P}$ is the rectangle P_0 in the statement of this theorem; thus P is I-special iff P_0 is $\tilde{I}$-critical.

Let us also observe that if $(P, Q, n) \in \mathcal{R}(\tilde{I})$ and P is I-special, i.e. $\tilde{I}$-critical, then it is $\tilde{I}$-special by Proposition 3 of Section 5.6.

The same discussion holds for Q, replacing P_0 by Q_k in the case where $(P, Q, n) \in \mathcal{R}(I) - \mathcal{R}(\tilde{I})$.

The following result will be useful in Section 7.

Proposition **11.** — *Let* $(P_0, Q_0, n_0), (P_1, Q_1, n_1) \in \mathcal{R}_N(I)$ *such that* $Q_0 \subset Q_u$, $P_1 \subset P_s$. *Assume that* $Q_0 \pitchfork_I P_1$ *holds and that* I, Q_0, P_1 *are maximal with this property: any* $\widehat{I} \supset I$, $(\widehat{P}_0, \widehat{Q}_0, \widehat{n}_0), (\widehat{P}_1, \widehat{Q}_1, \widehat{n}_1) \in \mathcal{R}_N(\widehat{I})$ *with* $Q_0 \subset \widehat{Q}_0 \subset Q_u$, $P_1 \subset \widehat{P}_1 \subset P_s$ *such that* $\widehat{Q}_0 \pitchfork_{\widehat{I}} \widehat{P}_1$ *holds must verify* $\widehat{I} = I$, $\widehat{Q}_0 = Q_0$, $\widehat{P}_1 = P_1$. *Then* Q_0 *and* P_1 *are* I-*special.*

Proof. — We may assume that the level of I is > 0. Let $\tilde{I}$ be the parent of I, and $(\tilde{P}_0, \tilde{Q}_0, \tilde{n}_0), (\tilde{P}_1, \tilde{Q}_1, \tilde{n}_1) \in \mathcal{R}_N(\tilde{I})$ be such that $\tilde{Q}_0$ (resp. $\tilde{P}_1$) is the smallest $\tilde{I}$-defined rectangle containing Q_0 (resp. P_1). We have to show that Q_0, P_1 are $\tilde{I}$-critical.

Assume for instance by contradiction that $\tilde{Q}_0$ is $\tilde{I}$-transverse. Let $(P_\alpha, Q_\alpha, n_\alpha)$ an $\tilde{I}$-decomposition of P_s such that, for each α, $\tilde{Q}_0$ and P_α are either $\tilde{I}$-separated or $\tilde{I}$-transverse. Let α such that $P_\alpha \cap P_1 \neq \emptyset$. Then $\tilde{Q}_0$ and P_α are $\tilde{I}$-transverse. If

$P_1 \subset P_\alpha$, it follows from Proposition 2 that $\widetilde{Q}_0$ and P_1 are $\widetilde{I}$-transverse. If $P_1 \supset P_\alpha$, then $(P_1, Q_1, n_1) \in \mathcal{R}_N(\widetilde{I})$ by coherence for $\mathcal{R}_N(I)$, and it follows then by Corollary 4 from $\widetilde{Q}_0 \pitchfork_{\widetilde{I}} P_\alpha$ and $Q_0 \pitchfork_I P_1$ that $\widetilde{Q}_0 \pitchfork_{\widetilde{I}} P_1$ holds.

Thus $\widetilde{Q}_0 \pitchfork_{\widetilde{I}} P_1$ always holds, in contradiction with the assumption of the lemma. This shows that $\widetilde{Q}_0$ is $\widetilde{I}$-critical. The proof that $\widetilde{P}_1$ is $\widetilde{I}$-critical is similar. □

6.6.2. *Uniform stretched exponential estimates 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

$$\gamma := \frac{\log \frac{3}{2}}{\log 2} \in (0, 1). \tag{6.45}$$

Proposition **12.** — *Let* (P, Q, n) *be an element of* $\mathcal{R}_N(I)$. *For all* $t \in I$, *we have*

$$|P| \leqslant C_0 \exp(-n^\gamma)$$

with the stronger estimate

$$|P| \leqslant C_0 \exp(-2n^\gamma)$$

when P *is* I-*special.*

The constant C_0 depends only on the constants in the formulas (3.12) and (3.27) for the widths in simple and parabolic composition.

Proof. — For $(P, Q, n) \in \mathcal{R}(I_0)$, we have an exponential estimate for $|P|$ which implies the weaker estimate of the I-special case of the proposition.

We assume now that the level of I is > 0.

For $(P, Q, n) \in \mathcal{R}(\widetilde{I}) - \mathcal{R}(I_0)$ (with $\widetilde{I}$ the parent of I), the result is true by the induction hypothesis (as observed in Section 6.6.1, if P is I-special, it is also $\widetilde{I}$-special).

We now assume that $(P, Q, n) \in \mathcal{R}(I) - \mathcal{R}(\widetilde{I})$ and apply the structure theorem of Section 6.5. We write

$$(P, Q, n) \in (P_0, Q_0, n_0) \,\square \cdots \square\, (P_k, Q_k, n_k) \tag{6.46}$$

as in the statement of the theorem. Let us denote by (P'_1, Q'_1, n'_1) the rectangle in $\mathcal{R}_{N-1}(I)$ defined by $(P_1, Q_1, n_1) \,\square \cdots \square\, (P_k, Q_k, n_k)$ such that Q is contained in Q'_1. As P_1 is $\widetilde{I}$-critical, P'_1 is I-special. We have

$$(P, Q, n) \in (P_0, Q_0, n_0) \,\square\, (P'_1, Q'_1, n'_1). \tag{6.47}$$

We use (3.27) (Section 3.5), condition (R7) for $\mathcal{R}_{N-1}(I)$ (Section 5.4) and the induction hypothesis to write

$$|P| \ll |P_0||P'_1|^{1/2} \leqslant C_0^{3/2} \exp(-n_0^\gamma - {n'_1}^\gamma). \tag{6.48}$$

As $n = n_0 + n'_1 + N_0$, this gives the first statement of the proposition.

Assume now that P is I-special. Then P_0 is $\widetilde{I}$-critical, hence $\widetilde{I}$-special (cf. Section 6.6.1).

When $n_0 \geqslant n'_1$, we have from the choice of γ that

$$(\textbf{6.49}) \qquad 2n_0^\gamma + {n'_1}^\gamma \geqslant 2(n_0 + n'_1)^\gamma,$$

and therefore the same estimate $|P| \ll |P_0||P'_1|^{1/2}$ now leads to $|P| \leqslant C_0 \exp(-2n^\gamma)$.

On the other hand, when $n_0 \leqslant n'_1$, we proceed as follows. We now have

$$(\textbf{6.50}) \qquad n_0^\gamma + 2{n'_1}^\gamma \geqslant 2(n_0 + n'_1)^\gamma.$$

As (P_0, Q_0, n_0) is $\widetilde{I}$-bicritical and $\widetilde{I}$ is regular, we have $|P_0| \leqslant |\widetilde{I}|^\beta$ for all $t \in \widetilde{I}$. From (3.27), we have

$$(\textbf{6.51}) \qquad |P_0| \leqslant C|P_0||P'_1|\delta(Q_0, P'_1)^{-1/2}.$$

From $Q_0 \pitchfork_I P'_1$, and (T1), (R7) (Section 5.4), we have $\delta(Q_0, P'_1) \geqslant |I|$ for all $t \in I$. Using also $|P_0| \leqslant |\widetilde{I}|^\beta$, we obtain

$$(\textbf{6.52}) \qquad |P| \ll |P_0|^{1/2}|P'_1| \leqslant C_0^{3/2} \exp(-n_0^\gamma - 2{n'_1}^\gamma)$$

and we conclude again that $|P| \leqslant C_0 \exp(-2n^\gamma)$ from (6.50). □

6.6.3. The following simple estimate says that the width of any rectangle not in $\mathcal{R}(I_0)$ is uniformly small with ε_0.

Proposition **13.** — *Let* $(P, Q, n) \in \mathcal{R}_N(I)$. *If* (P, Q, n) *does not belong to* $\mathcal{R}(I_0)$, *it satisfies, for all* $t \in I$

$$|P| < {\varepsilon_0}^{1/2}, \qquad |Q| < {\varepsilon_0}^{1/2}.$$

Proof. — We may assume that the level of I is > 0 and that (P, Q, n) does not belong to $\mathcal{R}(\widetilde{I})$, where $\widetilde{I}$ is the parent of I. Write

$$(\textbf{6.53}) \qquad (P, Q, n) \in (P_0, Q_0, n_0) \boxdot (P'_1, Q'_1, n'_1)$$

as in the proof of Proposition 12. The estimate (6.51) is still valid. From condition (R7) for $\mathcal{R}_{N-1}(I)$ (Section 5.4), we obtain $|P'_1| \ll \delta(Q_0, P'_1)$ and thus $|P| \ll |P_0|(\delta(Q_0, P'_1))^{1/2}$. As $\delta(Q_0, P'_1)$ is at most of the order of ε_0 and $|P_0|$ is small, we get the required estimate for P. The proof for $|Q|$ is symmetric. □

Lemma **3.** — *Let* (P_0, Q_0, n_0), (P_1, Q_1, n_1) *be elements of* $\mathcal{R}_{N-1}(I)$ *such that* $Q_0 \subset Q_u$ *and* $P_1 \subset P_s$. *If* $Q_0 \pitchfork_I P_1$ *holds then we have* $\delta(Q_0, P_1) > |I|$ *for all* $t \in I$.

Proof. — Let $\widetilde{\mathrm{I}} \supset \mathrm{I}$, and let elements $(\widetilde{\mathrm{P}}_0, \widetilde{\mathrm{Q}}_0, \widetilde{n}_0)$, $(\widetilde{\mathrm{P}}_1, \widetilde{\mathrm{Q}}_1, \widetilde{n}_1) \in \mathcal{R}(\widetilde{\mathrm{I}})$ with $\widetilde{\mathrm{P}}_1 \supset \mathrm{P}_1$, $\widetilde{\mathrm{Q}}_0 \supset \mathrm{Q}_0$ such that $\widetilde{\mathrm{Q}}_0 \,\overline{\pitchfork}_{\widetilde{\mathrm{I}}}\, \widetilde{\mathrm{P}}_1$ holds. From condition (T1), we have $\delta_{\mathrm{LR}}(\widetilde{\mathrm{Q}}_0, \widetilde{\mathrm{P}}_1) \geqslant 2|\widetilde{\mathrm{I}}|$ for all $t \in \mathrm{I}$. Using (3.36)–(3.39) and property (R7) for $\mathcal{R}_{\mathrm{N}-1}(\mathrm{I})$ we obtain, for all $t \in \mathrm{I}$

$$\delta(\mathrm{Q}_0, \mathrm{P}_1) \geqslant \delta(\widetilde{\mathrm{Q}}_0, \widetilde{\mathrm{P}}_1)$$

$$\geqslant \delta_{\mathrm{LR}}(\widetilde{\mathrm{Q}}_0, \widetilde{\mathrm{P}}_1) - \mathrm{C}(|\mathrm{Q}_0| + |\mathrm{P}_1|) > \frac{1}{2}\delta_{\mathrm{LR}}(\widetilde{\mathrm{Q}}_0, \widetilde{\mathrm{P}}_1) \geqslant |\widetilde{\mathrm{I}}| \geqslant |\mathrm{I}|. \tag{6.54}$$ □

In the following proposition, the setting and the notations are those of the structure theorem in Section 6.5: $\widetilde{\mathrm{I}}$ is the parent interval of I, (P, Q, n) is an element in $\mathcal{R}_{\mathrm{N}}(\mathrm{I})$ but not in $\mathcal{R}(\widetilde{\mathrm{I}})$, $k > 0$ and $(\mathrm{P}_i, \mathrm{Q}_i, n_i) \in \mathcal{R}(\widetilde{\mathrm{I}})$, $0 \leqslant i \leqslant k$ are as given by the conclusion of the theorem.

Proposition **14.** — *For all* $t \in \mathrm{I}$, *we have*

$$|\mathrm{P}| \leqslant \mathrm{C}^k |\mathrm{P}_0||\mathrm{P}_1| \cdots |\mathrm{P}_k||\mathrm{I}|^{-\frac{k}{2}}.$$

Moreover, for all $t \in \mathrm{I}$, *we have*

$$|\mathrm{P}_i| < |\widetilde{\mathrm{I}}|^{\beta} \quad \textit{for } 0 < i < k, \qquad |\mathrm{P}_k| < |\widetilde{\mathrm{I}}|.$$

In particular, we have always $|\mathrm{P}| < |\mathrm{P}_0||\mathrm{I}|^{1/3}$.

Proof. — For $0 \leqslant i \leqslant k$, let $(\mathrm{P}^{(i)}, \mathrm{Q}^{(i)}, p_i)$ be the element in $\mathcal{R}_{\mathrm{N}}(\mathrm{I})$ such that

$$(\mathrm{P}^{(i)}, \mathrm{Q}^{(i)}, p_i) \in (\mathrm{P}_0, \mathrm{Q}_0, n_0) \,\square \cdots \square\, (\mathrm{P}_i, \mathrm{Q}_i, n_i)$$

and $\mathrm{P} \subset \mathrm{P}^{(i)}$. We show by induction on i that

$$|\mathrm{P}^{(i)}| \leqslant \mathrm{C}^i |\mathrm{P}_0| \cdots |\mathrm{P}_i||\mathrm{I}|^{-\frac{i}{2}}. \tag{6.55}$$

As $\mathrm{P}^{(0)} = \mathrm{P}_0$, this is true for $i = 0$. By Lemma 3 above, we have, for all $t \in \mathrm{I}$,

$$\delta(\mathrm{Q}^{(i)}, \mathrm{P}_{i+1}) > |\mathrm{I}|,$$

so the required estimate for P_{i+1} follows from (3.27) in Section 3.5. This proves the first statement of the proposition.

The estimate $|\mathrm{P}_i| < |\widetilde{\mathrm{I}}|^{\beta}$ for $0 < i < k$ is true because $(\mathrm{P}_i, \mathrm{Q}_i, n_i)$ is $\widetilde{\mathrm{I}}$-bicritical and $\widetilde{\mathrm{I}}$ is β-regular. To estimate $|\mathrm{P}_k|$, we first observe that, by Proposition 3 in Section 5.6.1, there exists $t \in \widetilde{\mathrm{I}}$ such that $\delta_{\mathrm{LR}}(\mathrm{Q}_{k-1}, \mathrm{P}_k) < 2|\widetilde{\mathrm{I}}|$. Then, from Corollary 8 for $\mathcal{R}(\widetilde{\mathrm{I}})$ (in Section 7.6), we obtain, as $\mathrm{Q}_{k-1}, \mathrm{P}_k$ are $\widetilde{\mathrm{I}}$-critical, hence $\widetilde{\mathrm{I}}$-special, that $\delta_{\mathrm{LR}}(\mathrm{Q}_{k-1}, \mathrm{P}_k) < 4|\widetilde{\mathrm{I}}|$ for all $t \in \widetilde{\mathrm{I}}$. Finally, by property (R7) for $\mathcal{R}_{\mathrm{N}-1}(\mathrm{I})$, $|\mathrm{P}_k|$ is much smaller than $\delta_{\mathrm{LR}}(\mathrm{Q}_{k-1}, \mathrm{P}_k)$ for all $t \in \mathrm{I}$, which gives the required estimate.

The last statement in the proposition is an easy consequence of the first two. □

In the following corollary, the setting is the same that in the proposition. For $0 \leqslant i \leqslant k$, $(P^{(i)}, Q^{(i)}, p_i)$ is the element defined in the proof of the proposition.

Corollary **5.** — *Assume that* P *is* I*-special. Then, for all* $t \in I$, *we have*

$$|P^{(i)}| < |\widetilde{I}|^{\beta(i+1) - \frac{(1+\tau)i}{2}} \quad \text{for } 0 \leqslant i < k,$$
$$|P| < |\widetilde{I}|^{\beta k + 1 - \frac{(1+\tau)k}{2}}.$$

In particular, one has always $|P| < |I|^{\beta + 1/3}$.

Proof. — As P is I-special, P_0 is $\widetilde{I}$-critical. Therefore, (P_0, Q_0, n_0) is $\widetilde{I}$ -bicritical and $|P_0| < |\widetilde{I}|^{\beta}$ for all $t \in I$. The estimates of the corollary now follow easily from those of the proposition and (6.55) above. □

Corollary **6.** — *Let* $(P^*, Q^*, n^*) \in \mathcal{R}_N(I)$. *Assume that* P^* *is* I*-special and that* $|P^*| \geqslant |I|^{\beta+1/3}$ *for some* $t \in I$. *If* $|P^*| \geqslant \varepsilon_0^{\beta+1/3}$ *for some* $t \in I$, *let* $I^* = I_0$. *Otherwise, let* $I^* \supset I$ *be the smallest parameter interval such that* $|P^*| < |I^*|^{\beta+1/3}$ *for all* $t \in I$. *Then* $(P^*, Q^*, n^*) \in \mathcal{R}_N(I^*)$.

Proof. — If $I = I_0$, there is nothing to prove. Assume that the level of I is > 0, and let $\widetilde{I}$ be the parent interval. Corollary 5 above show that we must have $(P^*, Q^*, n^*) \in \mathcal{R}_N(\widetilde{I})$. Then P^* is $\widetilde{I}$-critical, hence $\widetilde{I}$-special. Iterating the argument gives the corollary. □

Observe that, once the existence of the class $\mathcal{R}(I)$ is established (for every candidate I in a β regular parent interval $\widetilde{I}$), the corollary implies that every candidate is $\overline{\beta}$-regular, with $\overline{\beta} = (1+\tau)^{-1}\beta$. Indeed, I-critical elements are either very thin ($\ll |I|^{\beta}$), or already $\widetilde{I}$-defined, and $|\widetilde{I}|^{\beta} = |I|^{\overline{\beta}}$.

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

7.1. *Uniform cone condition.* — In this subsection, we will check that all elements $(P, Q, n) \in \mathcal{R}_N(I)$ satisfy the cone condition (AL2) of Section 3.2 for the parameters λ, u_0, v_0 of Section 5.3: we have $u_0 = \frac{u}{(uv)^{1/4}}$, $v_0 = \frac{v}{(uv)^{1/4}}$, where all $(P, Q, n) \in \mathcal{R}(I_0)$ 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

$$\textbf{(AL2)} \qquad \lambda|A_x| + u_0|A_y| \leqslant 1,$$
$$\lambda|B_y| + v_0|B_x| \leqslant 1.$$

Let $u_1 = \frac{u}{(uv)^{1/8}}$, $v_1 = \frac{v}{(uv)^{1/8}}$. We will prove that, for all $t \in I$, we have

$$\textbf{(7.1)} \qquad |A_y| < u_1^{-1}, \qquad |B_x| < v_1^{-1}.$$

This is sufficient to obtain (AL2): we already know that if $(P, Q, n) \in \mathcal{R}(I_0)$ them (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 Proposition 13

$$(\mathbf{7.2}) \qquad |A_x| < \varepsilon_0^{\frac{1}{2}}, \qquad |B_y| < \varepsilon_0^{\frac{1}{2}}.$$

With ε_0 small enough, (7.1) and (7.2) give (AL2).

Let us now proceed with the proof of (7.1). When $(P, Q, n) \in \mathcal{R}(I_0)$, we have the stronger estimate:

$$(\mathbf{7.3}) \qquad |A_y| < u^{-1}, \qquad |B_x| < v^{-1}.$$

Assume that $(P, Q, n) \notin \mathcal{R}(I_0)$. By the structure theorem (Section 6.5), (P, Q, n) is obtained from the parabolic composition of shorter elements $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$, $(\widehat{P}, \widehat{Q}, \widehat{n}) \in \mathcal{R}_{N-1}(I)$. Denote by $(\widetilde{A}, \widetilde{B})$ the implicit representation of the affine-like iterate $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$. We use formula (A.86) of Appendix A to obtain

$$(\mathbf{7.4}) \qquad |A_y - \widetilde{A}_y| \leqslant C|\widetilde{P}||\widetilde{Q}|(\delta(\widetilde{Q}, \widehat{P}))^{-\frac{1}{2}}.$$

From (R7), $\delta(\widetilde{Q}, \widehat{P})$ is much larger than $|\widetilde{Q}|$. We have therefore

$$(\mathbf{7.5}) \qquad |A_y - \widetilde{A}_y| \leqslant |\widetilde{P}| \leqslant C_0 \exp(-\widetilde{n}^{\gamma}),$$

where we have used Proposition 12 in the last inequality. We only use (7.5) when $\widetilde{n}$ is large (because $\widetilde{Q} \subset Q_u$), and the series $\sum \exp(-m^{\gamma})$ is convergent. Therefore (7.1) is a consequence of (7.3) and (7.5).

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

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

$$(\mathbf{7.6}) \qquad D(g_t^n/P) \leqslant 2D_0.$$

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

$$(\mathbf{7.7}) \qquad D(g_t^n/P) \leqslant D_0.$$

Define

$$(\mathbf{7.8}) \qquad D_I(N) = \sup_{(P,Q,n)\in\mathcal{R}_N(I)} \sup_{t\in I} D(g_t^n/P).$$

From (7.7), we have $D_I(N) \leqslant D_0$ for $N = o(\log \varepsilon_0^{-1})$, because no parabolic composition is involved in this case. We set

$$D_I^s(N) := \max_{\substack{n>0, n'>0 \\ n+n' \leqslant N}} D_I^s(n, n') \tag{7.9}$$

with

$$D_I^s(n, n') := D_I(n) + C \exp(-n'^{\gamma})(D_I(n) + D_I(n')). \tag{7.10}$$

We also set

$$D_I^p(N) := \max_{\substack{n \gg 0, n' \gg 0 \\ n+n'+N_0 \leqslant N}} D_I^p(n, n') \tag{7.11}$$

with

$$D_I^p(n, n') := D_I(n) + C \exp(-\eta n'^{\gamma}). \tag{7.12}$$

We claim that, if $D_I(n)$ for $n < N$ is not too large so that the condition in (3.29) (Section 3.5) is satisfied, we have

$$D_I(N) \leqslant \max(D_I^s(N), D_I^p(N)). \tag{7.13}$$

Indeed, this follows from (3.13) (Section 3.3) for simple composition and (3.29) (Section 3.5) for parabolic composition; the term $C|P_1|\delta^{-1}$ in (3.29) is smaller than $|P_1|^{\eta}$ by condition (R7) for $\mathcal{R}_{N-1}(I)$; then one uses Proposition 12.

It is now clear that (7.6) follows from $D_I(N) \leqslant D_0$ for $N = o(\log \varepsilon_0^{-1})$ and (7.9)–(7.13).

7.3. *Estimates for the special rectangles* P_s *and* Q_u. — In Section 7.5, we will check the estimates contained in condition (R4) of Section 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

$$G_t : (R_{a_s}) \cap g_t^{-1}(R_{a_s}) \to g_t(R_{a_s}) \cap R_{a_s}. \tag{7.14}$$

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

As the equation of $W^s_{\mathrm{loc}}(p_s)$ is $\{x = 0\}$ (cf. (MP3) in Section 2.2), we have

$$A(y, 0, t) \equiv 0, \tag{7.15}$$

from which we deduce

$$\begin{aligned} |A_y(y, x, t)| &\leqslant C|x|, \\ |A_t(y, x, t)| &\leqslant C|x|, \\ |A_{yy}(y, x, t)| &\leqslant C|x|, \\ |A_{yt}(y, x, t)| &\leqslant C|x|. \end{aligned} \tag{7.16}$$

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

$$C^{-1}\mu^{-n} \leqslant |A_x^{(n)}(y, x, t)| \leqslant C\mu^{-n}. \tag{7.17}$$

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

$$C^{-1}\mu^{m-\ell}|x_\ell| \leqslant |x_m| \leqslant C\mu^{m-\ell}|x_\ell|. \tag{7.18}$$

Proposition **15.** — *The following estimates hold:*

$$|A_y^{(n)}(y_0, x_n, t)| \leqslant C|x_0| \leqslant C\mu^{-n}|x_n|, \tag{7.19}$$

$$|A_t^{(n)}(y_0, x_n, t)| \leqslant Cn|x_0| \leqslant Cn\mu^{-n}|x_n|, \tag{7.20}$$

$$|A_{yy}^{(n)}(y_0, x_n, t)| \leqslant C|x_0| \leqslant C\mu^{-n}|x_n|. \tag{7.21}$$

Proof of (7.19). — From formula (3.11) in Section 3.3, we have:

$$A_y^{(n)}(y_0, x_n, t) = A_y^{(n-1)}(y_0, x_{n-1}, t) + A_y A_x^{(n-1)} B_y^{(n-1)} \Delta^{-1}, \tag{7.22}$$

with $B_y^{(n-1)}\Delta^{-1}$ exponentially small with n and, using (7.16)–(7.18):

$$|\, A_y(y_{n-1}, x_n, t) A_x^{(n-1)}(y_0, x_{n-1}, t)\,| \leqslant C|x_0|. \tag{7.23}$$

The inequality (7.19) is now clear. □

Proof of (7.20). — We use here formulas (A.6), (A.10) of Appendix A which give

$$\begin{aligned} &|\, A_t^{(n)}(y_0, x_n, t) - A_t^{(n-1)}(y_0, x_{n-1}, t)\,| \\ &\qquad \leqslant C\mu^{-n}\Big(|\, A_t(y_{n-1}, x_n, t)\,| + |\, B_t^{(n-1)}(y_0, x_{n-1}, t)\,||\, A_y(y_{n-1}, x, t)\,|\Big), \end{aligned} \tag{7.24}$$

$$|B_t^{(n)}(y_0, x_n, t) - B_t^{(n-1)}(y_1, x_n, t)| \leqslant C|B_y^{(n-1)}|_{C^0}\Big(1 + |A_t^{(n-1)}|_{C^0}\Big). \tag{7.25}$$

As $|B_y^{(n)}|_{C_0}$ is exponentially small, we deduce from (7.25) that

$$|B_t^{(n)}(y_0, x_n, t)| \leqslant C, \tag{7.26}$$

and then, from (7.24), (7.16) that (7.20) holds. □

Proof of (7.21). — We use formulas (A.6), (A.11), (A.18), (A.20) of Appendix A to obtain

$$A_{yy}^{(n)} = A_{yy}^{(n-1)} + 2A_{xy}^{(n-1)}X_y + A_{xx}^{(n-1)}X_y^2 + A_x^{(n-1)}X_{yy}, \tag{7.27}$$

with

$$X_y = A_y B_y^{(n-1)}\Delta^{-1}, \tag{7.28}$$

$$\Delta = 1 - A_y B_x^{(n-1)}, \tag{7.29}$$

$$\begin{aligned} X_{yy} = B_y^{(n-1)}\Delta^{-1}\Big(&A_{yy}B_y^{(n-1)}\Delta^{-1} + A_y\partial_y \log|B_y^{(n-1)}| \\ &+ A_y X_y \partial_x \log|B_y^{(n-1)}| - A_y\Delta_y\Delta^{-1}\Big), \end{aligned} \tag{7.30}$$

$$-\Delta_y = A_{yy}B_y^{(n-1)}B_x^{(n-1)}\Delta^{-1} + A_y B_{xy}^{(n-1)} + A_y B_{xx}^{(n-1)}X_y. \tag{7.31}$$

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 $B_y^{(n-1)}$, $B_{xy}^{(n-1)}$, $A_x^{(n-1)}$, $A_{xx}^{(n-1)}$, $A_{xy}^{(n-1)}$ are exponentially small. Also, from (7.16) we have:

$$\begin{aligned} &|A_y(y_{n-1}, x_n, t)| \leqslant C|x_n|, \\ &|A_{yy}(y_{n-1}, x_n, t)| \leqslant C|x_n|. \end{aligned} \tag{7.32}$$

We conclude that we can write

$$A_{yy}^{(n)}(y_0, x_n, t) = A_{yy}^{(n)}(y_0, x_{n-1}, t) + \mu^{-n}x_n r_n \tag{7.33}$$

with r_n exponentially small; this leads to (7.21). □

Corollary **7.** — *For the special rectangle* (P_s, Q_s, n_s), *we have:*

$$|A_y^{(n_s)}|_{C^0} \leqslant C\varepsilon_0,$$

$$|A_{yy}^{(n_s)}|_{C^0} \leqslant C\varepsilon_0,$$

$$|A_t^{(n_s)}|_{C^0} \leqslant C\varepsilon_0 \log \varepsilon_0^{-1}.$$

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).

7.4. *Further estimates for the class* $\mathcal{R}(I_0)$. — In this subsection, we derive estimates for the class $\mathcal{R}(I_0)$ from the estimates for (P_s, Q_s, n_s), (P_u, Q_u, n_u) obtained in the last subsection and from the estimates for simple composition found in Appendix A.3. These estimates are typically better than the estimates for classes $\mathcal{R}(I)$ with smaller I, which will be obtained in the end of Section 7, because only simple composition is involved into the construction of elements of $\mathcal{R}(I_0)$.

Proposition **16.** — *Let* (P, Q, n) *be an element of* $\mathcal{R}(I_0)$*, and let* (A, B) *be the associated implicit representation. For all* t, y, x*, we have*

$$|A_t| \leqslant C, \qquad |B_t| \leqslant C,$$
$$|A_{yt}| \leqslant C, \qquad |B_{xt}| \leqslant C,$$
$$|\partial_t \log |A_x|| \leqslant Cn, \qquad |\partial_t \log |B_y|| \leqslant Cn.$$

If moreover $P \subset P_s$*, we have, for all* t, y, x

$$|A_y| \leqslant C\varepsilon_0, \qquad |A_{yy}| \leqslant C\varepsilon_0, \qquad |A_t| \leqslant C\varepsilon_0 |\log \varepsilon_0|.$$

Similarly, if $Q \subset Q_u$*, we have, for all* t, y, x

$$|B_x| \leqslant C\varepsilon_0, \qquad |B_{xx}| \leqslant C\varepsilon_0, \qquad |B_t| \leqslant C\varepsilon_0 |\log \varepsilon_0|.$$

Proof. — The widths of rectangles in $\mathcal{R}(I_0)$ are exponentially small with the number of iterations. Then, the estimate for A_t, B_t follow by a simple induction on n from formula (A.67) in Appendix A.3. Similarly, one derives the estimates for A_{yt}, B_{xt}, $\partial_t \log |A_x|$, $\partial_t \log |B_y|$ from formulas (A.72), (A.73).

When $P \subset P_s$, we write

$$(P, Q, n) = (P_s, Q_s, n_s) * (P', Q', n')$$

for some $(P', Q', n') \in \mathcal{R}(I_0)$. As $|P_s| \leqslant C\varepsilon_0$, the estimates for A_y, A_{yy}, A_t follow from Corollary 7 and formulas (A.66), (A.71), (A.67) in Appendix A.3. The case where $Q \subset Q_u$ is similar. □

7.5. *Proof of the property* (R4) *of Section 5.3.* — We have to show that, for an element (P, Q, n) in $\mathcal{R}_N(I)$, with associated implicit representation (A, B), we have

$$|A_y| \leqslant C\varepsilon_0, \qquad |A_{yy}| \leqslant C\varepsilon_0 \tag{7.34}$$

whenever $P \subset P_s$ and

$$(\mathbf{7.35}) \qquad |B_x| \leqslant C\varepsilon_0, \qquad |B_{xx}| \leqslant C\varepsilon_0$$

whenever $Q \subset Q_u$. We will deal only with (7.34), the other case being symmetric.

We have already proved (7.34), (7.35) when (P, Q, n) belongs to $\mathcal{R}(I_0)$. We may therefore assume that (P, Q, n) does not belong to $\mathcal{R}(I_0)$. We assume that $P \subset P_s$ and prove (7.34).

In this case, the structure theorem of Section 6.5 guarantees that (P, Q, n) can be obtained from the parabolic composition of shorter elements (P_0, Q_0, n_0), (P_1, Q_1, n_1). Let (A_0, B_0), (A_1, B_1) be the implicit representations associated to these iterates.

From formulas (A.86), (A.91) in Appendix A.4, we have

$$|A_y - A_{0,y}| \leqslant C|P_0||Q_0|\delta(Q_0, P_1)^{-1/2},$$
$$|A_{yy} - A_{0,yy}| \leqslant C|P_0||Q_0|\delta(Q_0, P_1)^{-1/2}.$$

As $P \subset P_s$, we have also $P_0 \subset P_s$, hence $|P_0| \leqslant C\varepsilon_0$. On the other hand, by property (R7) for $\mathcal{R}_{N-1}(I)$, $|Q_0|$ is for all $t \in I$ much smaller than $\delta(Q_0, P_1)$. We have then, from Proposition 12 in Section 6.6.2

$$(\mathbf{7.36}) \qquad |Q_0|\delta(Q_0, P_1)^{-1/2} \leqslant |Q_0|^{1/2} \leqslant C \exp\left(-\frac{1}{2}n_0^{\gamma}\right),$$

which gives

$$(\mathbf{7.37}) \qquad |A_y - A_{0,y}| \leqslant C\varepsilon_0, \exp\left(-\frac{1}{2}n_0^{\gamma}\right),$$

$$(\mathbf{7.38}) \qquad |A_{yy} - A_{0,yy}| \leqslant C\varepsilon_0 \exp\left(-\frac{1}{2}n_0^{\gamma}\right).$$

As $P_0 \subset P_s$, the estimate (7.34) follows immediately by induction on n (starting with Proposition 16) from (7.37), (7.38). The proof of (7.35) is similar.

7.6. *Relative speeds of special rectangles.* — Let (P_0, Q_0, n_0), (P_1, Q_1, n_1) be elements of $\mathcal{R}_N(I)$ 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 (3.30)–(3.33) of Section 3.6 and are the values at the four corners of the rectangle of definition of the function $\overline{C}(y_0, x_1)$ introduced in Section 3.5 as

$$(\mathbf{7.39}) \qquad \overline{C}(y_0, x_1) = \min_w C(w, y_0, x_1).$$

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 obtain good estimates when Q_0, P_1 are I-special, which is sufficient in the applications.

Proposition **17.** — *Let* (P, Q, n) *be an element of* $\mathcal{R}_N(I)$ *with* $P \subset P_s$. *Let* (A, B) *be the implicit representation of* (P, Q, n). *If* P *is* I-*special, we have for all* t, y, x

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

Proof. — If $(P, Q, n) \in \mathcal{R}(I_0)$, we have from Proposition 14 the stronger estimate:

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

We now assume that the level of I is > 0. Let $\widetilde{I}$ be the parent of I. If $(P, Q, n) \in \mathcal{R}(\widetilde{I})$, P is also $\widetilde{I}$-special (cf. Section 6.6.1) and the estimate of the proposition is true by induction.

We now assume that $(P, Q, n) \notin \mathcal{R}(\widetilde{I})$. We apply the structure theorem in Section 6.5: let $k \geqslant 1$ and (P_i, Q_i, n_i), $0 \leqslant i \leqslant k$, be the elements of $\mathcal{R}(\widetilde{I})$ given by the statement of the theorem. We also denote, for $0 \leqslant i \leqslant k$, by $(P^{(i)}, Q^{(i)}, n^{(i)})$ the element of $\mathcal{R}(\widetilde{I})$ such that $P \subset P^{(i)}$ and

$$(P^{(i)}, Q^{(i)}, n^{(i)}) \in (P_0, Q_0, n_0) \boxdot \cdots \boxdot (P_i, Q_i, n_i). \tag{7.40}$$

We have $(P^{(k)}, Q^{(k)}, n^{(k)}) = (P, Q, n)$. As P is I-special, P_0 is $\widetilde{I}$-critical, hence $P^{(i)}$ is I-special for $0 \leqslant i \leqslant k$. Moreover, for $0 \leqslant i < k$, Q_i, P_{i+1} are $\widetilde{I}$-critical, hence $Q^{(i)}$, P_{i+1} are I-special.

Let $(A^{(i)}, B^{(i)})$, (A_i, B_i) be the implicit representations associated to $(P^{(i)}, Q^{(i)}, n^{(i)})$, (P_i, Q_i, n_i). The estimate (A.96) in Appendix A.4 gives

$$|A_t^{(i+1)} - A_t^{(i)}| \leqslant C|P^{(i)}|\delta(Q^{(i)}, P_{i+1})^{-1/2}(1 + |A_{i+1,t}| + |B_t^{(i)}|). \tag{7.41}$$

Here we have from the induction hypothesis $|B_t^{(i)}| \leqslant C\varepsilon_0$ since $Q^{(i)}$ is I-special and $|A_{i+1,t}| \leqslant C\varepsilon_0$ since P_{i+1} is I-special.

From Lemma 3 in Section 6.6.3 we have $\delta(Q^{(i)}, P_{i+1}) > |I|$ for all $t \in I$. From Corollary 5 in Section 6.6.3, we have

$$|P^{(i)}| \leqslant C^i |\widetilde{I}|^{\beta(i+1) - \frac{(1+\tau)i}{2}}. \tag{7.42}$$

Plugging these estimates into (7.41) above gives

$$|A_t^{(i+1)} - A_t^{(i)}| \leqslant C^{i+1} |\widetilde{I}|^{(i+1)(\beta - \frac{1+\tau}{2})}. \tag{7.43}$$

As $\beta - \frac{1+\tau}{2} > \frac{1}{2}$, we obtained the required estimate for $|A_t|$ by first summing over i and then on the successive levels of the parameter intervals under consideration.

The proof for B_t is similar. □

Corollary **8.** — *Let* (P_0, Q_0, n_0), (P_1, Q_1, n_1) *be elements of* $\mathcal{R}_N(I)$ *with* $Q_0 \subset Q_u$, $P_1 \subset P_s$. *Assume that* Q_0 *and* P_1 *are* I*-special. Then, the partial derivative with respect to the parameter of the function* C *(introduced in Section 3.5) satisfies*

$$|C_t + 1| \leqslant C\varepsilon_0^{\frac{1}{2}}.$$

Proof. — From formula (A.35) in Appendix A, using the notations there, we have

$$-C_t = \theta_x \overline{X}_t + \theta_y \overline{Y}_t + \theta_t, \tag{7.44}$$

$$\overline{X}_t = (A_{1,t} + A_{1,y} Y_{s,t})\Delta_1^{-1}, \tag{7.45}$$

$$\overline{Y}_t = (B_{0,t} + B_{0,x} X_{u,t})\Delta_0^{-1}, \tag{7.46}$$

with Δ_0^{-1}, Δ_1^{-1} uniformly bounded. The value of θ_t is taken at $(\overline{X}, \overline{Y}, t)$, with

$$|\overline{X}| \leqslant C\varepsilon_0, \qquad |\overline{Y}| \leqslant C\varepsilon_0. \tag{7.47}$$

On the other hand, we have, in Section 4.1, normalized the parameter in order to have

$$\theta_t(0, 0, t) \equiv 1. \tag{7.48}$$

We, therefore, have

$$|\theta_t(\overline{X}, \overline{Y}, t) - 1| \leqslant C\varepsilon_0. \tag{7.49}$$

In (7.45) and (7.46), 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 17. The Corollary follows, as θ_x, θ_y, $Y_{s,t}$, $X_{u,t}$ are uniformly bounded. □

Corollary **9.** — *Let* (P_0, Q_0, n_0), (P_1, Q_1, n_1) *be elements of* $\mathcal{R}_N(I)$ *with* $Q_0 \subset Q_u$, $P_1 \subset P_s$. *Assume that* Q_0 *and* P_1 *are* I*-special and that* $Q_0 \,\overline{\pitchfork}_I\, P_1$ *holds. Then, we have*

$$\max_I \delta_{LR}(Q_0, P_1) \leqslant 2 \min_I \delta_{LR}(Q_0, P_1).$$

Proof. — Indeed, by condition (T1) of Section 5.4, we have $\min_I \delta_{LR}(Q_0, P_1) \geqslant 2|I|$ and by Corollary 8 above, we have $\max_I \delta_{LR}(Q_0, P_1) - \min_I \delta_{LR}(Q_0, P_1) < 2|I|$. □

We will have a more general version of this statement in the next Subsection (Proposition 19).

7.7. *Variation of width of special rectangles.* — Our main purpose now is to prove property (R7) of Section 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 $Q_0 \pitchfork_I P_1$ holds, then, for all $t \in I$, we have

$$\delta(Q_0, P_1) \geqslant C^{-1}\Big(|P_1|^{1-\eta} + |Q_0|^{1-\eta}\Big).$$

A priori, the transversality condition gives some control through (T2), (T3) in Section 5.4 only for *some* values of the parameter. However, from Corollary 9 above, we know that the order of magnitude of $\delta(Q_0, P_1)$ is the same through I, at least when Q_0, P_1 are I-special and $Q_0 \overline{\pitchfork}_I P_1$ holds. 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 when Q_0, P_1 are I-special, and this will turn out to be sufficient due to Proposition 11 in Section 6.6.1.

Proposition **18.** — *Let* (P, Q, n) *be an element of* $\mathcal{R}_N(I)$ *with* $P \subset P_s$, (A, B) *be the associated implicit representation. If* P *is* I-*special, we have, for all* t, y, x

$$|\partial_t \log|A_x|| \leqslant C\frac{\log|P|}{|I|\log|I|}, \tag{7.50}$$

$$|A_{yt}| \leqslant C. \tag{7.51}$$

Obviously, there is a similar statement exchanging P and Q, A and B, x and y.

Proof. — 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 of Section 7.4 imply the inequalities above.

We will now assume that the level of I is > 0. Let $\tilde{I}$ be the parent of I. If $(P, Q, n) \in \mathcal{R}(\tilde{I})$, P is also $\tilde{I}$-special (cf. Section 6.6.1) and $|I||\log|I|| < |\tilde{I}||\log|\tilde{I}||$. Therefore the estimates of the proposition follow from the induction hypothesis.

We now assume that $(P, Q, n) \notin \mathcal{R}(\tilde{I})$. We apply the structure theorem in Section 6.5: let $k \geqslant 1$ and (P_i, Q_i, n_i), $0 \leqslant i \leqslant k$, be the elements of $\mathcal{R}(\tilde{I})$ given by the statement of the theorem. We also denote, for $0 \leqslant i \leqslant k$, by $(P^{(i)}, Q^{(i)}, n^{(i)})$ the element of $\mathcal{R}(\tilde{I})$ such that $P \subset P^{(i)}$ and

$$(P^{(i)}, Q^{(i)}, n^{(i)}) \in (P_0, Q_0, n_0) \boxdot \cdots \boxdot (P_i, Q_i, n_i). \tag{7.52}$$

We have $(P^{(k)}, Q^{(k)}, n^{(k)}) = (P, Q, n)$. As P is I-special, P_0 is $\tilde{I}$-critical, hence $P^{(i)}$ is I-special for $0 \leqslant i \leqslant k$. Moreover, for $0 \leqslant i < k$, Q_i, P_{i+1} are $\tilde{I}$-critical, hence $Q^{(i)}$, P_{i+1} are I-special.

Let $(A^{(i)}, B^{(i)})$, (A_i, B_i) be the implicit representations associated to $(P^{(i)}, Q^{(i)}, n^{(i)})$, (P_i, Q_i, n_i). The estimates (A.103), (A.104) in Appendix A.4 give, with $\delta = \delta(Q^{(i)}, P_{i+1})$

$$|\partial_t \log|A_x^{(i+1)}| - \partial_t \log|A_x^{(i)}||$$
$$\leqslant C(\delta^{-1} + \delta^{-\frac{1}{2}}(|B_{xt}^{(i)}| + |A_{i+1,yt}|) + |\partial_t \log|A_{i+1,x}||), \tag{7.53}$$

$$|A_{yt}^{(i+1)} - A_{yt}^{(i)}| \leqslant C\delta^{-\frac{1}{2}}|P^{(i)}|(1 + |Q^{(i)}|K), \tag{7.54}$$

with

$$K = \delta^{-1} + \delta^{-\frac{1}{2}}(|B_{xt}^{(i)}| + |A_{i+1,yt}|) + |\partial_t \log |A_x^{(i)}|| + |\partial_t \log |B_y^{(i)}||. \tag{7.55}$$

Assume, by the induction hypothesis, that we have

$$|B_{xt}^{(i)}| \leqslant C_0, \tag{7.56}$$

$$|A_{i+1,yt}| \leqslant C_0, \tag{7.57}$$

$$|\partial_t \log |A_{i+1,x}|| \leqslant C_0 \frac{\log |P_{i+1}|}{|\widetilde{I}| \log |\widetilde{I}|}, \tag{7.58}$$

$$|\partial_t \log |A_x^{(i)}|| \leqslant C_0 \frac{\log |P^{(i)}|}{|I| \log |I|}, \tag{7.59}$$

$$|\partial_t \log |B_y^{(i)}|| \leqslant C_0 \frac{\log |Q^{(i)}|}{|I| \log |I|}. \tag{7.60}$$

Here C_0 is large but independent of ε_0. This means that the term $\delta^{-\frac{1}{2}}(|B_{xt}^{(i)}| + |A_{i+1,yt}|)$ in (7.53) and (7.55) is dominated by δ^{-1}. As $|I| = |\widetilde{I}|^{1+\tau}$, in order to prove (7.50) by induction, we need to have, in view of (7.53):

$$C|I||\log |I||\delta^{-1} + CC_0|\widetilde{I}|^{\tau}|\log |P_{i+1}|| + C_0|\log |P^{(i)}|| \leqslant C_0|\log |P^{(i+1)}||. \tag{7.61}$$

We have here $\delta \geqslant |I|$ from Lemma 3 in Section 6.6.3 and, by (3.27):

$$|\log |P^{(i+1)}|| \geqslant |\log |P^{(i)}|| + |\log |P_{i+1}|| - \frac{1}{2}|\log |I|| - C. \tag{7.62}$$

Therefore, (7.61) will hold as far as

$$\left(\frac{C_0}{2} + C\right)|\log |I|| + C_0 C \leqslant C_0(1 - C|\widetilde{I}|^{\tau})|\log |P_{i+1}||. \tag{7.63}$$

From (R7) for $\mathcal{R}_{N-1}(I)$, we know that $|P_{i+1}|$ is much smaller than δ. On the other hand, as Q_j and P_{i+1} are not $\widetilde{I}$-transverse, by Proposition 3 in Section 5.6.1, $\delta_{LR}(Q_j, P_{i+1})$ (which is larger than δ) is smaller than $2|\widetilde{I}|$ for some $t \in \widetilde{I}$; it then follows from Corollary 8 in the last subsection applied to Q_j, P_{i+1} (which are $\widetilde{I}$-critical, hence $\widetilde{I}$-special), that $\delta_{LR}(Q_j, P_{i+1})$, and thus also δ, stay smaller than $C|\widetilde{I}|$ for all $t \in \widetilde{I}$. We therefore have

$$|\log |P_{i+1}|| \geqslant \log |\widetilde{I}| = (1 + \tau) \log |I|, \tag{7.64}$$

from which (7.63) follows if we take $C_0 \geqslant 3C$ (provided ε_0 is small enough). This completes the proof of the induction step for (7.50).

To prove (7.51), we estimate the right-hand side of (7.54). From Corollary 5 in Section 6.6.3, we have

$$|P^{(i)}| \leqslant C^i |\widetilde{I}|^{\beta(i+1) - \frac{(1+\tau)i}{2}}. \tag{7.65}$$

From (R7) for $\mathcal{R}_{N-1}(I)$, we know that $|Q^{(i)}|$ is much smaller than δ. We have seen above that $|I| < \delta < C|\widetilde{I}|$.

This gives, as $\beta > 1$

$$\delta^{-\frac{1}{2}} |P^{(i)}| \leqslant |\widetilde{I}|^{(i+1)\beta/2}, \tag{7.66}$$

$$\delta^{-\frac{3}{2}} |P^{(i)}||Q^{(i)}| \leqslant |\widetilde{I}|^{(i+1)\beta/2}, \tag{7.67}$$

$$\delta^{-\frac{1}{2}} |P^{(i)}||Q^{(i)}||\partial_t \log |A_x^{(i)}|| \leqslant |\widetilde{I}|^{(i+1)\beta/2}, \tag{7.68}$$

$$\delta^{-\frac{1}{2}} |P^{(i)}||Q^{(i)}||\partial_t \log |B_y^{(i)}|| \leqslant |\widetilde{I}|^{(i+1)\beta/2}. \tag{7.69}$$

This leads to:

$$|A_{yt}^{(i+1)} - A_{yt}^{(i)}| \leqslant C|\widetilde{I}|^{\beta/2(i+1)}. \tag{7.70}$$

We can now sum over i and then over the different levels of parameter intervals to obtain (7.51). The proof of Proposition 18 is complete. □

Proof of Property (R7). — Let (P_0, Q_0, n_0), (P_1, Q_1, n_1) be elements of $\mathcal{R}_N(I)$ such that $Q_0 \subset Q_u$, $P_1 \subset P_s$ and $Q_0 \pitchfork_I P_1$ holds.

If $\widehat{I} \supset I$, $(\widehat{P}_0, \widehat{Q}_0, \widehat{n}_0)$, $(\widehat{P}_1, \widehat{Q}_1, \widehat{n}_1) \in \mathcal{R}_N(\widehat{I})$ satisfy $Q_0 \subset \widehat{Q}_0 \subset Q_u$, $P_1 \subset \widehat{P}_1 \subset P_s$, and the inequality in (R7) is satisfied for $\widehat{I}$, $\widehat{Q}_0$, $\widehat{P}_1$ then it is also satisfied for I, Q_0, P_1.

We can therefore assume that I, Q_0, P_1 are maximal with the property $Q_0 \pitchfork_I P_1$. From Proposition 11 in Section 6.6.1, Q_0, P_1 are then I-special.

The maximality property implies that $Q_0 \overline{\pitchfork}_I P_1$ holds. Then, from Corollary 9 in Section 7.6, we have

$$\max_I \delta_{LR}(Q_0, P_1) \leqslant 2 \min_I \delta_{LR}(Q_0, P_1). \tag{7.71}$$

We have also $\min_I \delta_{LR}(Q_0, P_1) \geqslant 2|I|$ from (T1) for $Q_0 \overline{\pitchfork}_I P_1$.

From Proposition 18, we have either $\max_I |Q_0| < |I|^2$ or $\max_I |Q_0| \leqslant C \min_I |Q_0|$, and similarly for P_1. It then follows from (T2), (T3) for $Q_0 \overline{\pitchfork}_I P_1$ that, for all $t \in I$

$$\delta(Q_0, P_1) \geqslant C^{-1}(|Q_0|^{1-\eta} + |P_1|^{1-\eta})$$

for all $t \in I$.

The proof of property (R7) is complete. □

Once (R7) has been obtained, we have a stronger form of Corollary 9 in Section 7.6:

Proposition **19.** — *Let* (P_0, Q_0, n_0), (P_1, Q_1, n_1) *be elements of* $\mathcal{R}_N(I)$ *with* $Q_0 \subset Q_u$, $P_1 \subset P_s$. *Assume that* $Q_0 \pitchfork_I P_1$ *holds. Then, we have*

$$\max_I \delta_{LR}(Q_0, P_1) \leqslant 3 \min_I \delta(Q_0, P_1).$$

Proof. — If $\widehat{I} \supset I$, $(\widehat{P}_0, \widehat{Q}_0, \widehat{n}_0)$, $(\widehat{P}_1, \widehat{Q}_1, \widehat{n}_1) \in \mathcal{R}_N(\widehat{I})$ satisfy $Q_0 \subset \widehat{Q}_0 \subset Q_u$, $P_1 \subset \widehat{P}_1 \subset P_s$, and the inequality in the Proposition is satisfied for $\widehat{I}, \widehat{Q}_0, \widehat{P}_1$ then it is also satisfied for I, Q_0, P_1.

We can therefore assume that I, Q_0, P_1 are maximal with the property $Q_0 \pitchfork_I P_1$. From Proposition 11 in Section 6.6.1, Q_0, P_1 are then I-special. The maximality property also implies that $Q_0 \,\overline{\pitchfork}_I\, P_1$ holds. By Corollary 9 and (R7), we have then

$$\begin{aligned}\max_I \delta_{LR}(Q_0, P_1) &\leqslant 2 \min_I \delta_{LR}(Q_0, P_1)\\ &\leqslant 2 \min_I (\delta(Q_0, P_1) + C|Q_0| + C|P_1|)\\ &\leqslant 3 \min_I \delta(Q_0, P_1).\end{aligned}$$

□

The existence and properties of the classes $\mathcal{R}(I)$, for a candidate parameter interval I are now fully justified and the proof of the theorem in Section 5.4 is complete. What we do not know at this moment is whether there exists, besides I_0, 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, 10 and 11.

8. Number of children and dimension estimates

In this section (except in Section 8.4), we fix a candidate parameter interval I (in particular, the parent interval $\widetilde{I}$ is always assumed to be β-regular) and obtain some properties of the class $\mathcal{R}(I)$ which will be important in Sections 9 and 10.

In Section 8.1, we collect some results that could have been proved earlier, but whose proof was deferred because the results were not necessary for the construction of the class $\mathcal{R}(I)$.

In Section 8.2, we establish some bounds on the number of children of a given rectangle which will be useful both in Sections 8.3, 8.4 and in Sections 9, 10 and 11.

In Section 8.3, we prove that, in some appropriate sense, the transverse stable and unstable "dimensions" of the class $\mathcal{R}(I)$ are very close to those of $\mathcal{R}(I_0)$.

Finally, in Section 8.4, we establish an estimate in parameter space which will be essential in Section 9.

8.1. *Further criteria for transversality.* — The following results are useful variants of Proposition 10 in Section 6.4.

Proposition **20.** — *Let* (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.*

There is a symmetric statement exchanging P's and Q's.

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 concavity. □

Proposition **21.** — *Let* (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* $Q_0 \subset Q_u$, $P_1 \subset P'_1 \subset P_s$. *Assume that* $Q_0 \pitchfork_I P_1$ *holds and that* $|P'_1| \leqslant \frac{1}{2}|Q_0|$ *for all* $t \in I$. *Then* Q_0 *and* P'_1 *are also* I*-transverse.*

Again, here is a symmetric statement exchanging P's and Q's.

Proof. — This follows closely the proof of Proposition 10. By definition of the transversality relation, there exist $\widetilde{I} \supset I$, $(\widetilde{P}_0, \widetilde{Q}_0, \widetilde{n}_0)$, $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1) \in \mathcal{R}(\widetilde{I})$ such that $Q_0 \subset \widetilde{Q}_0$, $P_1 \subset \widetilde{P}_1$ and $\widetilde{Q}_0 \,\overline{\pitchfork}_{\widetilde{I}}\, \widetilde{P}_1$.

If $P'_1 \subset \widetilde{P}_1$ this already implies that $Q_0 \pitchfork_I P'_1$. Let us assume that $\widetilde{P}_1 \subset P'_1$. We will show that $\widetilde{Q}_0 \,\overline{\pitchfork}_{\widetilde{I}}\, P'_1$ holds. By coherence (Proposition 8), we have $(P'_1, Q'_1, n'_1) \in \mathcal{R}(\widetilde{I})$.

Let us check (T1)–(T3).

By (T1) for $Q_0 \pitchfork_I P_1$ and (3.49) in Section 3.5, we have, for all $t \in \widetilde{I}$:

$$\delta_{LR}(\widetilde{Q}_0, P'_1) \geqslant \delta_{LR}(\widetilde{Q}_0, \widetilde{P}_1) \geqslant 2|\widetilde{I}|. \tag{8.1}$$

By (T2) for $Q_0 \pitchfork_I P_1$ and (3.47), there exists $t_0 \in \widetilde{I}$ such that

$$\delta_R(\widetilde{Q}_0, P'_1) \geqslant \delta_R(\widetilde{Q}_0, \widetilde{P}_1) \geqslant 2|\widetilde{Q}_0|^{1-\eta}. \tag{8.2}$$

Finally, for this same value t_0, we have

$$\begin{aligned} \delta_L(\widetilde{Q}_0, P'_1) &\geqslant \delta_R(\widetilde{Q}_0, P'_1) - C|P'_1| \\ &\geqslant 2|Q_0|^{1-\eta} - C|P'_1| \\ &\geqslant 2(2|P'_1|)^{1-\eta} - C|P'_1| \\ &\geqslant 2|P'_1|^{1-\eta}, \end{aligned} \tag{8.3}$$

if ε_0 is small enough. This proves that $\widetilde{Q}_0 \,\overline{\pitchfork}_{\widetilde{I}}\, P'_1$ holds and thus also $Q_0 \pitchfork_I P'_1$ holds. □

Corollary **10.** — *Let* $(P_1, Q_1, n_1) \in \mathcal{R}(I)$. *Assume that* P_1 *is* I*-critical. Then there exists* $(P_0, Q_0, n_0) \in \mathcal{R}(I)$ *such that* Q_0 *and* P_1 *are* I*-critically related and* $|Q_0| < \max(|I|^\beta, 2|P_1|)$ *for some* $t \in I$.

Proof. — Let $(P, Q, n) \in \mathcal{R}(I)$ such that Q and P_1 are I-critically related and $|Q| \geqslant 2|P_1|$ for all $t \in I$. We claim that Q is I-critical. Indeed, assume on the contrary that Q is I-transverse. Let (P_i, Q_i, n_i) be an I-decomposition of P_s such that, for each i, Q is either I-separated from P_i or I-transverse to P_i. If P_1 is contained in some P_i, Q is I-separated from P_1 if it is I-separated from P_i, and I-transverse to P_1 if it is I-transverse to P_i. If P_1 contains some P_i such that Q is I-transverse to P_i, then Q is I-transverse to P_1 by the proposition above. The remaining case is where the P_i contained in P_1 form an I-decomposition of P_1 and Q is I-separated from each of them; then Q is I-separated from P_1. In all cases, we get a contradiction. This proves the claim.

Consider now the following inductively constructed sequence of I-decompositions of Q_u, starting with the canonical one by children of Q_u. We stop the process when one element $(P_\alpha, Q_\alpha, n_\alpha)$ at least in the decomposition is such that $Q_\alpha \pitchfork_I P_1$ does not hold and $|Q_\alpha| < \max(|I|^\beta, 2|P_1|)$ for some $t \in I$: it satisfies then the conclusion of the Corollary.

To go from one I-decomposition to the next one, we keep those $(P_\alpha, Q_\alpha, n_\alpha)$ such that Q_α is I-transverse to P_1, or I-separated from P_1. Because P_1 is I-critical, there are other elements $(P_\alpha, Q_\alpha, n_\alpha)$ in the decomposition, which are I-critically related to P_1 and satisfy $|Q_\alpha| \geqslant \max(|I|^\beta, 2|P_1|)$ for all $t \in I$ by assumption. The claim above shows that Q_α is I-critical and therefore I-decomposable (Corollary 3 in Section 5.6.5). We replace such a Q_α by its children.

It is clear that the process has to stop, and the proof of the Corollary is complete. □

Proposition **22.** — *Let* $I' \supset I$ *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* $Q_0 \subset Q_u$, $P_1 \subset P_s$. *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.*

Again, there is a symmetric statement exchanging P's and Q's.

Proof. — Let $(\widehat{P}_0, \widehat{Q}_0, \widehat{n}_0), (\widehat{P}_1, \widehat{Q}_1, \widehat{n}_1) \in \mathcal{R}(I')$ satisfy $Q_0 \subset \widehat{Q}_0 \subset Q_u$, $P_1 \subset \widehat{P}_1 \subset P_s$; then $2|I'| < |\widehat{P}_1|^{1-\eta}$ for all $t \in I'$, and if $\widehat{Q}_0$ and $\widehat{P}_1$ are I'-transverse, then Q_0 and P_1 are I'-transverse.

It is therefore sufficient to consider the case where Q_0, P_1, I are maximal with the properties $Q_0 \pitchfork_I P_1, I \subset I'$.

If $I = I'$, we are done. We therefore assume that I is strictly smaller than I'. In this case, by Proposition 11 in Section 6.6.1, Q_0 and P_1 are I-special. Maximality also guarantees that $Q_0 \overline{\pitchfork}_I P_1$.

We show that $Q_0 \overline{\pitchfork}_{I'} P_1$. Properties (T2), (T3) for $Q_0 \overline{\pitchfork}_I P_1$ imply the same for $Q_0 \overline{\pitchfork}_{I'} P_1$.

For the value t_1 given by (T3), we have

$$\delta_{\mathrm{LR}}(\mathrm{Q}_0, \mathrm{P}_1) \geqslant \delta_{\mathrm{L}}(\mathrm{Q}_0, \mathrm{P}_1) \geqslant 2|\mathrm{P}_1|^{1-\eta} \geqslant 4|\mathrm{I}'|. \tag{8.4}$$

By Corollary 8 in Section 7.6, this implies

$$\delta_{\mathrm{LR}}(\mathrm{Q}_0, \mathrm{P}_1) \geqslant 2|\mathrm{I}'|, \quad \forall t \in \mathrm{I}',$$

which is (T1) for $\mathrm{Q}_0 \overline{\pitchfork}_{\mathrm{I}'} \mathrm{P}_1$. This contradicts the maximality of I and proves the proposition. □

8.2. *Estimates on the number of children.* — We start with some preliminary results.

Proposition **23.** — *Let* $\mathrm{I}' \supset \mathrm{I}$ *be a parameter interval, and let* $(\widetilde{\mathrm{P}}, \widetilde{\mathrm{Q}}, \widetilde{n})$ *be an element of* $\mathcal{R}(\mathrm{I}')$. *We assume that* $\widetilde{\mathrm{Q}}$ *is* I'*-transverse. Then, any element* $(\mathrm{P}, \mathrm{Q}, n)$ *in* $\mathcal{R}(\mathrm{I})$ *such that* P *is a child of* $\widetilde{\mathrm{P}}$ *is already an element of* $\mathcal{R}(\mathrm{I}')$.

Proof. — We can assume that P is a non-simple child. Then $(\mathrm{P}, \mathrm{Q}, n)$ is obtained by parabolic composition in $\mathcal{R}(\mathrm{I})$ of $(\widetilde{\mathrm{P}}, \widetilde{\mathrm{Q}}, \widetilde{n})$ with some $(\mathrm{P}_1, \mathrm{Q}_1, n_1) \in \mathcal{R}(\mathrm{I})$. As $\widetilde{\mathrm{Q}}$ is I'-transverse, there exists an I'-decomposition $(\mathrm{P}_\alpha, \mathrm{Q}_\alpha, n_\alpha)$ of P_s such that each P_α is I'-separated or I'-transverse with $\widetilde{\mathrm{Q}}$. Let α_0 be such that P_{α_0} and P_1 intersect. Then, $\widetilde{\mathrm{Q}} \pitchfork_{\mathrm{I}'} \mathrm{P}_{\alpha_0}$ holds, and also $\widetilde{\mathrm{Q}} \pitchfork_{\mathrm{I}} \mathrm{P}_1$; if we had $\mathrm{P}_1 \subsetneqq \mathrm{P}_{\alpha_0}$, this would imply that $\widetilde{\mathrm{Q}}$ is I-transverse to the parent $\widetilde{\mathrm{P}}_1$ of P_1 and P would not be a child of $\widetilde{\mathrm{P}}$. Therefore, we must have $\mathrm{P}_{\alpha_0} \subset \mathrm{P}_1$. By coherence (Proposition 8), we have that $(\mathrm{P}_1, \mathrm{Q}_1, n_1) \in \mathcal{R}(\mathrm{I}')$. By concavity (Proposition 9), from $\widetilde{\mathrm{Q}} \pitchfork_{\mathrm{I}} \mathrm{P}_1$ and $\widetilde{\mathrm{Q}} \pitchfork_{\mathrm{I}'} \mathrm{P}_{\alpha_0}$, we deduce that $\widetilde{\mathrm{Q}} \pitchfork_{\mathrm{I}'} \mathrm{P}_1$ also holds and $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}(\mathrm{I}')$. □

Proposition **24.** — *Let* $\mathrm{I}_0 \supset \mathrm{I}_1 \supset \mathrm{I}$ *be the largest parameter interval such that*

$$|\mathrm{I}_1|^\beta < \left(\frac{1}{2}|\mathrm{I}|\right)^{\frac{1}{1-\eta}}. \tag{8.5}$$

Let $(\widetilde{\mathrm{P}}, \widetilde{\mathrm{Q}}, \widetilde{n})$, $(\mathrm{P}, \mathrm{Q}, n)$ *be elements of* $\mathcal{R}(\mathrm{I})$ *such that* P *is a non-simple child of* $\widetilde{\mathrm{P}}$. *Let* $(\mathrm{P}_1, \mathrm{Q}_1, n_1)$, $(\widetilde{\mathrm{P}}_1, \widetilde{\mathrm{Q}}_1, \widetilde{n}_1)$ *be the elements of* $\mathcal{R}(\mathrm{I})$ *such that*

$$(\mathrm{P}, \mathrm{Q}, n) \in (\widetilde{\mathrm{P}}, \widetilde{\mathrm{Q}}, \widetilde{n}) \,\square\, (\mathrm{P}_1, \mathrm{Q}_1, n_1)$$

and $\widetilde{\mathrm{P}}_1$ *is the parent of* P_1.

Then, $(\mathrm{P}_1, \mathrm{Q}_1, n_1)$ *belongs to* $\mathcal{R}(\mathrm{I}_1)$, $\widetilde{\mathrm{P}}_1$ *is* I*-critical,* $\widetilde{\mathrm{Q}}_1$ *is* I_1*-transverse and we have*

$$2|\widetilde{\mathrm{P}}_1|^{1-\eta} > |\mathrm{I}| \tag{8.6}$$

for all $t \in \mathrm{I}$.

Remark **9.** — As parabolic composition is possible, we have $I \neq I_0$; then, as $\beta > (1-\eta)^{-1}$, we must have $I_1 \supsetneqq I$ and I_1 is β-regular.

Proof. — As P is a child of $\widetilde{P}$, $\widetilde{Q} \pitchfork_I \widetilde{P}_1$ does not hold. Then, it follows from Proposition 19 above that $\widetilde{P}_1$ is I-critical, and from Proposition 11 that (8.6) holds. Then, by definition of I_1, we have:

$$|\widetilde{P}_1| > |I_1|^{\beta} \tag{8.7}$$

for all $t \in I$. Let us show that $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1)$ belongs to $\mathcal{R}(I_1)$. Otherwise, there would exist $I_2 \supset I$, with parent interval $\widetilde{I}_2 \subset I_1$, such that $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1)$ belongs to $\mathcal{R}(I_2)$ but not to $\mathcal{R}(\widetilde{I}_2)$. Then, we would have from Corollary 5 in Section 6.6.3 (as $\widetilde{P}_1$ is I-critical, hence I_2-special) that $|\widetilde{P}_1| < |I_2|^{\beta+1/3}$ for all $t \in I_2$, in contradiction to (8.7).

Therefore, $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1)$ belongs to $\mathcal{R}(I_1)$. As $I_1 \supsetneqq I$, I_1 is β-regular; $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1)$ cannot be I_1-bicritical in view of (8.7); $\widetilde{P}_1$ is I_1-critical and hence $\widetilde{Q}_1$ is I_1-transverse. Proposition 23 then shows that $(P_1, Q_1, n_1) \in \mathcal{R}(I_1)$. □

Corollary **11.** — *Assume that the level of* I *is* > 0, *and let* $\widetilde{I}$ *be the parent interval of* I. *Let* $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$ *be an element of* $\mathcal{R}(\widetilde{I})$, *such that* $|\widetilde{Q}|^{1-\eta} > 2|I|$ *for all* $t \in \widetilde{I}$. *Then all children of* $\widetilde{P}$ *in* $\mathcal{R}(I)$ *belong already to* $\mathcal{R}(\widetilde{I})$.

Proof. — Indeed, let $(P, Q, n) \in \mathcal{R}(I)$ such that P is a child of $\widetilde{P}$. We can assume that P is a non-simple child. Let I_1, (P_1, Q_1, n_1), $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1) \in \mathcal{R}(I_1)$ be as in Proposition 24. We have $I_1 \supset \widetilde{I}$. As $\widetilde{Q} \pitchfork_I P_1$, we can apply Proposition 21 in Section 8.1 to obtain that $\widetilde{Q} \pitchfork_{\widetilde{I}} P_1$ holds and $(P, Q, n) \in \mathcal{R}(\widetilde{I})$. □

Corollary **12.** — *Let* $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$ *be an element of* $\mathcal{R}(I)$. *The number of* $(P, Q, n) \in \mathcal{R}(I)$ *such that* P *is a child of* $\widetilde{P}$ *is finite.*

Proof. — We argue by induction on the level of the parameter interval I.

If I is the starting interval I_0, $\widetilde{P}$ has only simple children and the assertion is obvious. Assume that $I \subsetneqq I_0$. 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 $\widetilde{P}$, let I_1, (P_1, Q_1, n_1), $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1) \in \mathcal{R}(I_1)$ be as in Proposition 24. By the induction hypothesis, there is for each fixed $\widetilde{P}_1$ only a finite number of possibilities for P_1. On the other hand, in view of relation (8.6), there are obviously only a finite number of possibilities for $\widetilde{P}_1$. The induction step is complete, and this completes the proof. □

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 $\widetilde{P}_1$, and the number of possibilities for P_1 once $\widetilde{P}_1$ is fixed.

Proposition **25.** — *Let* $(\widetilde{\mathrm{P}}, \widetilde{\mathrm{Q}}, \widetilde{n})$ *be an element of* $\mathcal{R}(\mathrm{I})$. *The number of* $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}(\mathrm{I})$ *such that* P *is a child of* $\widetilde{\mathrm{P}}$ *is at most* $|\mathrm{I}|^{-c\eta}$.

The constant c, as the other constants denoted by C, and the content c' in the next proposition, depends only on the initial diffeomorphism f, not on $\tau \gg \eta \gg \varepsilon_0$.

Proof. — We argue again by induction on the level of I, following the proof of Corollary 12. The notations are those of Proposition 24. When $\mathrm{I} = \mathrm{I}_0$, 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 $\mathrm{I} \neq \mathrm{I}_0$, the number of possibilities for P_1 when $\widetilde{\mathrm{P}}_1$ is fixed in at most $|\mathrm{I}_1|^{-c\eta}$ by the induction hypothesis. We have to estimate the number of possibilities for $\widetilde{\mathrm{P}}_1$. We know that $\widetilde{\mathrm{Q}} \pitchfork_{\mathrm{I}} \widetilde{\mathrm{P}}_1$ does not hold, but $\widetilde{\mathrm{Q}} \pitchfork_{\mathrm{I}} \mathrm{P}_1$ holds.

As $\widetilde{\mathrm{P}}_1$ is I-critical, it is I-special. We have from (8.6) and Proposition 18 in Section 7.7

$$(\mathbf{8.8}) \qquad \max_{\mathrm{I}} |\widetilde{\mathrm{P}}_1| \leqslant \mathrm{C} \min_{\mathrm{I}} |\widetilde{\mathrm{P}}_1|.$$

From Proposition 19 in Section 7.7, we have also, using (8.8)

$$(\mathbf{8.9}) \qquad \max_{\mathrm{I}} \delta(\widetilde{\mathrm{Q}}, \widetilde{\mathrm{P}}_1) < \max_{\mathrm{I}} \delta(\widetilde{\mathrm{Q}}, \mathrm{P}_1) \leqslant \mathrm{C} \min_{\mathrm{I}} \delta(\widetilde{\mathrm{Q}}, \mathrm{P}_1) < \mathrm{C} \min_{\mathrm{I}} (\delta(\widetilde{\mathrm{Q}}, \widetilde{\mathrm{P}}_1) + |\widetilde{\mathrm{P}}_1|).$$

Lemma **4.** — *We have for all* $t \in \mathrm{I}$

$$(\mathbf{8.10}) \qquad \delta(\widetilde{\mathrm{Q}}, \widetilde{\mathrm{P}}_1) \leqslant \mathrm{C} |\widetilde{\mathrm{P}}_1|^{1-\eta}.$$

Proof. — As $\widetilde{\mathrm{Q}} \overline{\pitchfork}_{\mathrm{I}} \widetilde{\mathrm{P}}_1$ does not hold, at least one of the following three properties must be true:

$$(\mathbf{8.11}) \qquad \delta_{\mathrm{LR}}(\widetilde{\mathrm{Q}}, \widetilde{\mathrm{P}}_1) < 2|\mathrm{I}| \quad \text{for some } t_0 \in \mathrm{I};$$

$$(\mathbf{8.12}) \qquad \delta_{\mathrm{R}}(\widetilde{\mathrm{Q}}, \widetilde{\mathrm{P}}_1) < 2|\widetilde{\mathrm{Q}}|^{1-\eta} \quad \text{for all } t \in \mathrm{I};$$

$$(\mathbf{8.13}) \qquad \delta_{\mathrm{L}}(\widetilde{\mathrm{Q}}, \widetilde{\mathrm{P}}_1) < 2|\widetilde{\mathrm{P}}_1|^{1-\eta} \quad \text{for all } t \in \mathrm{I}.$$

If (8.13) holds, we are done.

If (8.12) holds, we argue as follows. As $\widetilde{\mathrm{Q}} \pitchfork_{\mathrm{I}} \mathrm{P}_1$ holds but $\widetilde{\mathrm{Q}} \pitchfork_{\mathrm{I}} \widetilde{\mathrm{P}}_1$ does not hold, it follows from Proposition 20 in Section 8.1 that $|\widetilde{\mathrm{P}}_1| > \frac{1}{2}|\widetilde{\mathrm{Q}}|$ for some $t_0 \in \mathrm{I}$. Then, from (8.8) and (8.9), we have

$$\begin{aligned} \max_{\mathrm{I}} \delta(\widetilde{\mathrm{Q}}, \widetilde{\mathrm{P}}_1) &\leqslant \mathrm{C} \min_{\mathrm{I}} (\delta(\widetilde{\mathrm{Q}}, \widetilde{\mathrm{P}}_1) + |\widetilde{\mathrm{P}}_1|) \leqslant \mathrm{C} \min_{\mathrm{I}} (|\widetilde{\mathrm{Q}}|^{1-\eta} + |\widetilde{\mathrm{P}}_1|) \\ &\leqslant \mathrm{C} \max_{\mathrm{I}} |\widetilde{\mathrm{P}}_1|^{1-\eta} \leqslant \mathrm{C} \min_{\mathrm{I}} |\widetilde{\mathrm{P}}_1|^{1-\eta}. \end{aligned}$$

Finally, assume that (8.11) holds. As $\widetilde{Q} \pitchfork_I P_1$ holds but $\widetilde{Q} \pitchfork_I \widetilde{P}_1$ does not hold, it follows from Proposition 9 in Section 6.4 that $2|\widetilde{P}_1|^{1-\eta} > |I|$ for all $t \in I$. Then, from (8.8) and (8.9), we have

$$\max_I \delta(\widetilde{Q}, \widetilde{P}_1) \leqslant C \min_I (\delta(\widetilde{Q}, \widetilde{P}_1) + |\widetilde{P}_1|) \leqslant C(|I| + \max_I |\widetilde{P}_1|)$$
$$\leqslant C \max_I |\widetilde{P}_1|^{1-\eta} \leqslant C \min_I |\widetilde{P}_1|^{1-\eta}. \qquad \square$$

We are now able to estimate the number of possibilities for $\widetilde{P}_1$ and show that this number is at most

$$\textbf{(8.14)} \qquad C|I|^{-\frac{\eta}{1-\eta}}.$$

This indeed follows from (8.9), (8.6) and the fact that if two distinct $\widetilde{P}_1$ are not disjoint, the ratio of their widths is bounded away from 1 (so the $\widetilde{P}_1$ at a given scale are disjoint; one then sums over scales). The total number of children is thus bounded by

$$\textbf{(8.15)} \qquad C + 2C|I|^{-\frac{\eta}{1-\eta}} |I_1|^{-c\eta},$$

where I_1 was the largest parameter interval satisfying

$$\textbf{(8.16)} \qquad |I_1|^{\beta} < \left(\frac{1}{2}|I|\right)^{\frac{1}{1-\eta}}.$$

If $|I| > 2\varepsilon_0^{\beta(1-\eta)}$, we have $I_1 = I_0$; in this case, the term $|I_1|^{-c\eta}$ in (8.15) is unnecessary because $\widetilde{P}_1$ has only simple children. If $|I| \leqslant 2\varepsilon_0^{\beta(1-\eta)}$, we have

$$\textbf{(8.17)} \qquad |I_1|^{\beta(1+\tau)^{-1}} \geqslant \left(\frac{1}{2}|I|\right)^{\frac{1}{1-\eta}}$$

and the term in (8.15) is bounded by $|I|^{-c\eta}$ provided that ε_0 is small enough and

$$\textbf{(8.18)} \qquad c\eta > \frac{\eta}{1-\eta} + c\eta \frac{1+\tau}{1-\eta} \beta^{-1}.$$

As η, τ are very small, any choice of $c > \frac{\beta}{\beta-1}$ yields (8.18). Then, as $c > 1$, such a choice is also convenient when $|I| > 2\varepsilon_0^{\beta(1-\eta)}$, and this concludes the proof of Proposition 25. $\square$

In Proposition 25, we have estimated the total number of children in terms of the level of the parameter interval.

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 **26.** — *Let* $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$ *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* $\widetilde{P}$ *satisfying* $|P| \geqslant \varepsilon|\widetilde{P}|$ *for some* $t \in I$, *is at most* $\varepsilon^{-c'\eta}$.

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 $\widetilde{P}$. We assume, for some $t_0 \in I$, that:

$$|P| \geqslant \varepsilon|\widetilde{P}|. \tag{8.19}$$

Let (P_1, Q_1, n_1), $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1) \in \mathcal{R}(I)$ be as in Proposition 24. From (3.27), we have, for all $t \in I$:

$$|P| \leqslant C|\widetilde{P}||P_1|\delta(\widetilde{Q}, P_1)^{-\frac{1}{2}}. \tag{8.20}$$

Property (R7) guarantees that, for all $t \in I$

$$\delta(\widetilde{Q}, P_1) \geqslant C^{-1}|P_1|^{1-\eta}. \tag{8.21}$$

Combining (8.19), (8.20), (8.21), we have, for some $t_0 \in I$

$$\delta(\widetilde{Q}, P_1) \geqslant C^{-1}\varepsilon^{2\frac{1-\eta}{1+\eta}}. \tag{8.22}$$

As we always have

$$\delta(\widetilde{Q}, P_1) \leqslant C\varepsilon_0, \tag{8.23}$$

there is no non-simple child satisfying (8.19) unless $\varepsilon < \varepsilon_0^{\frac{1}{2}}$; we will assume that this holds in the sequel.

From Lemma 4 above, we have, for all $t \in I$:

$$\delta(\widetilde{Q}, \widetilde{P}_1) \leqslant C|\widetilde{P}_1|^{1-\eta}, \tag{8.24}$$

and thus, from (3.36), also

$$\begin{aligned} \delta(\widetilde{Q}, P_1) &\leqslant \delta_R(\widetilde{Q}, \widetilde{P}_1) \\ &\leqslant \delta(\widetilde{Q}, \widetilde{P}_1) + C|\widetilde{P}_1| \\ &\leqslant C|\widetilde{P}_1|^{1-\eta}. \end{aligned} \tag{8.25}$$

Combining (8.22) and (8.25), we get, for some $t_0 \in I$

$$|\widetilde{P}_1| \geqslant C^{-1}\varepsilon^{\frac{2}{1+\eta}}, \tag{8.26}$$

an inequality which actually holds for all $t \in I$ in view of (8.8).

As in the proof of Proposition 25, the number of $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1)$ for which both (8.25), (8.26) hold is easily seen to be at most

$$\textbf{(8.27)} \qquad C\varepsilon^{-\frac{2\eta}{1+\eta}}.$$

To estimate the number of P_1 for a given $\widetilde{P}_1$, we will apply Proposition 25 in an appropriate way.

Define a parameter interval $\widehat{I} \supset I$ as follows. If $|I|^\beta \geqslant \varepsilon^2$, let $\widehat{I} = I$. If $|I|^\beta < \varepsilon^2$, let $\widehat{I} \supset I$ be the largest parameter interval such that $|\widehat{I}|^\beta < \varepsilon^2$. We have thus in any case

$$\textbf{(8.28)} \qquad |\widehat{I}| \geqslant \min(\varepsilon_0, \varepsilon^{\frac{2(1+\tau)}{\beta}}).$$

As $\widetilde{P}_1$ is I-critical, hence I-special, it follows from (8.26) and Corollary 6 in Section 6.6.3 that $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1)$ belongs to $\mathcal{R}(\widehat{I})$. Moreover, when $\widehat{I} \neq I$, we have $|\widetilde{P}_1| > |\widehat{I}|^\beta$ for all $t \in I$ from (8.26) and the definition of $\widehat{I}$. Therefore, $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1)$ cannot be $\widehat{I}$-bicritical. As $\widetilde{P}_1$ is $\widehat{I}$-critical, $\widetilde{Q}_1$ is $\widehat{I}$-transverse and we conclude from Proposition 23 that (P_1, Q_1, n_1) also belongs to $\mathcal{R}(\widehat{I})$. The same is also obviously true when $\widehat{I} = I$.

We apply Proposition 25: for each fixed $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1)$, the number of children P_1 is at most $|\widehat{I}|^{-c\eta}$.

From (8.28), we have, for some appropriate c_0

$$\textbf{(8.29)} \qquad |\widehat{I}|^{-c\eta} \leqslant \varepsilon^{-c_0\eta}.$$

Combining this with the previous estimate in (8.27) for the number of possibilities for $\widetilde{P}_1$ gives therefore the required estimate. □

8.3. *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

$$\textbf{(8.30)} \qquad \Theta(P^*, I, s) = \sum \max_I |P|^s,$$

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 dependence of the estimate on P^* is also quite straightforward, through the simple scaling factor $\max_I |P^*|^s$.

Let us recall that we denote by d_s^0 the transverse Hausdorff dimension of the stable foliation $W^s(K)$ of the horseshoe K for the value 0 of the parameter. Let us also denote by d_s^t the transverse Hausdorff dimension of the stable foliation $W^s(K_{g_t})$ for the value t of the parameter. It is well-known that d_s^t depends smoothly on the parameter. The transverse Hausdorff dimension 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 any $t \in I_0$ and all $\varepsilon > 0$ the number of $(P, Q, n) \in \mathcal{R}(I_0)$ such that $|P| \geqslant \varepsilon$ is exactly of the order of $\varepsilon^{-d_s^t}$.

Let us say that P (with $(P, Q, n) \in \mathcal{R}(I)$) is a *simple* descendant of P^* if (P, Q, n) is the simple composition of (P^*, Q^*, n^*) by an element of $\mathcal{R}(I_0)$. Consider the series

$$\Theta_0(P^*, s) = \sum \max_I |P|^s, \tag{8.31}$$

where the sum runs over simple descendants of P^*. By what we have just recalled, for any $t \in I_0$, the series converge in a half-plane $\{\operatorname{Re} s > d_s^0 + C\varepsilon_0\}$ and satisfies, for $s \geqslant d_s^0 + 2C\varepsilon_0$

$$|\Theta_0(P^*, s)| < C \max_I |P^*|^s (\operatorname{Re} s - d_s^0)^{-1}. \tag{8.32}$$

For $\Theta(P^*, I, s)$ with a parameter interval $I \neq I_0$, we have to allow a slightly larger margin with relation to the initial value d_s^0.

Proposition **27.** — *Let* $(P^*, Q^*, n^*) \in \mathcal{R}(I)$. *The series* $\Theta(P^*, I, s)$ *is uniformly convergent in the half-plane* $\{\operatorname{Re} s \geqslant d_s^0 + \varepsilon_0^{\frac{1}{3}d_s^0}\}$ *and we have, for* $s \geqslant d_s^0 + \varepsilon_0^{\frac{1}{5}d_s^0}$

$$|\,\Theta(P^*, I, s) - \Theta_0(P^*, s)\,| \leqslant \max_I |P^*|^s \varepsilon_0^{\frac{1}{20}d_s^0}.$$

Proof. — Let (P, Q, n) be an element of $\mathcal{R}(I)$ with $P \subset P^*$. Consider the intermediary rectangles

$$P^* = P(0) \subset P(1) \subset \cdots \subset P(\ell) = P$$

with $P(i)$ the parent of $P(i+1)$. Let

$$\ell_0 < \ell_1 < \cdots < \ell_{k-1} \tag{8.33}$$

be the indices such that $P(\ell_j + 1)$ is a non-simple child of $P(\ell_j)$. Obviously, P is a simple descendant of P^* iff $k = 0$.

We also define for $0 \leqslant j \leqslant k$ elements $(P^{(j)}, Q^{(j)}, n^{(j)}) \in \mathcal{R}(I_0)$ by the following properties

$$(P(\ell_0), Q(\ell_0), n(\ell_0)) = (P(0), Q(0), n(0)) * (P^{(0)}, Q^{(0)}, n^{(0)}), \tag{8.34}$$

$$\begin{aligned}(P(\ell_j), Q(\ell_j), n(\ell_j)) = {} & (P(\ell_{j-1}+1), Q(\ell_{j-1}+1), n(\ell_{j-1}+1)) \\ & * (P^{(j)}, Q^{(j)}, n^{(j)}),\end{aligned} \tag{8.35}$$

$$(\mathrm{P},\mathrm{Q},n) = (\mathrm{P}(\ell_{k-1}+1),\mathrm{Q}(\ell_{k-1}+1),n(\ell_{k-1}+1)) \\ * (\mathrm{P}^{(k)},\mathrm{Q}^{(k)},n^{(k)}). \tag{8.36}$$

We now estimate the widths from (3.12), for all $t \in \mathrm{I}$:

$$|\mathrm{P}(\ell_0)| \leqslant \mathrm{C}|\mathrm{P}^*||\mathrm{P}^{(0)}|, \tag{8.37}$$

$$|\mathrm{P}(\ell_j)| \leqslant \mathrm{C}|\mathrm{P}(\ell_{j-1}+1)||\mathrm{P}^{(j)}|, \tag{8.38}$$

$$|\mathrm{P}| \leqslant \mathrm{C}|\mathrm{P}(\ell_{k-1}+1)||\mathrm{P}^{(k)}|. \tag{8.39}$$

From (3.27) and property (R7), we also have:

$$|\mathrm{P}(\ell_j+1)| < \varepsilon_0^{\frac{1}{2}}|\mathrm{P}(\ell_j)|. \tag{8.40}$$

Define m_j for $0 \leqslant j < k$ to be the largest integer such that, for all $t \in \mathrm{I}$

$$|\mathrm{P}(\ell_j+1)| \leqslant 2^{-m_j}\varepsilon_0^{\frac{1}{2}}|\mathrm{P}(\ell_j)|. \tag{8.41}$$

From Proposition 26, for each fixed $\mathrm{P}(\ell_j)$, the number of non-simple children $\mathrm{P}(\ell_j+1)$ satisfying (8.41) is at most

$$\left(2^{m_j+1}\varepsilon_0^{-\frac{1}{2}}\right)^{c'\eta}. \tag{8.42}$$

Combining (8.37), (8.38), (8.39) and (8.41), we also have

$$\max_{\mathrm{I}}|\mathrm{P}| \leqslant \mathrm{C}^{k+1}\max_{\mathrm{I}}|\mathrm{P}^*|\left(\prod_0^k \max_{\mathrm{I}}|\mathrm{P}^{(j)}|\right)\varepsilon_0^{\frac{k}{2}}2^{-\Sigma_0^{k-1}m_j}, \tag{8.43}$$

with the usual convention that $\Sigma_0^{k-1}m_j = 0$ when $k = 0$.

We will take this to the power s and sum over P. We introduce (corresponding to the term $|\mathrm{P}^{(j)}|^s$)

$$\Theta_0(s) := \sum_{(\widehat{\mathrm{P}},\widehat{\mathrm{Q}},\widehat{n})\in\mathcal{R}(\mathrm{I}_0)} \max_{\mathrm{I}}|\widehat{\mathrm{P}}|^s \\ = \sum_a \Theta(\mathrm{R}_a,\mathrm{I},s), \tag{8.44}$$

and also

$$\theta(s) := \sum_{m\geqslant 0}(\mathrm{C}\varepsilon_0^{\frac{1}{2}}2^{-m})^s(2^{m+1}\varepsilon_0^{-\frac{1}{2}})^{c'\eta}. \tag{8.45}$$

The function Θ_0 is controlled by (8.32), while θ satisfies

$$\theta(s) = 2^{c'\eta}\mathrm{C}^s\varepsilon_0^{\frac{1}{2}(s-c'\eta)}\left(1-2^{-(s-c'\eta)}\right)^{-1}, \tag{8.46}$$

and therefore, for $C^{-1} < s < C$:

$$C^{-1}\varepsilon_0^{\frac{1}{2}(s-c'\eta)} \leqslant \theta(s) \leqslant C\varepsilon_0^{\frac{1}{2}(s-c'\eta)}. \tag{8.47}$$

From (8.32), we have, for $s > d_s^0 + 2C\varepsilon_0$:

$$\Theta_0(s) \leqslant C(s - d_s^0)^{-1}. \tag{8.48}$$

In particular, for $s \geqslant d_s^0 + \varepsilon_0^{1/3d_s^0}$, we have

$$\Theta_0(s) \leqslant C\varepsilon_0^{-\frac{1}{3}d_s^0}, \tag{8.49}$$

$$\Theta_0(s)\theta(s) \leqslant C\varepsilon_0^{\frac{1}{10}d_s^0}. \tag{8.50}$$

But, from (8.43), we have for real s

$$\Theta(P^*, I, s) - \Theta_0(P^*, s) \leqslant C^s \max_I |P^*|^s \sum_{k>0} \Theta_0^{k+1}(s)\theta^k(s), \tag{8.51}$$

and therefore we deduce from (8.50) that the series defining Θ is uniformly convergent in the half plane $\{\mathrm{Re}\, s \geqslant d_s^0 + \varepsilon_0^{\frac{1}{3}d_s^0}\}$.

For $s > d_s^0 + \varepsilon_0^{\frac{1}{5}d_s^0}$, we have, from (8.48), (8.47):

$$\Theta_0(s) \leqslant C\varepsilon_0^{-\frac{1}{5}d_s^0}, \tag{8.52}$$

$$\Theta_0^2\theta(s) \leqslant C\varepsilon_0^{\frac{1}{15}d_s^0}, \tag{8.53}$$

which gives the second part of the proposition. □

Corollary **13.** — *Let* $d_s^* = d_s^0 + \varepsilon_0^{\frac{1}{5}d_s^0}$, $(P^*, Q^*, n^*) \in \mathcal{R}(I)$, $\varepsilon > 0$. *The number of* $(P, Q, n) \in \mathcal{R}(I)$ *with* $P \subset P^*$, $|P| > \varepsilon|P^*|$ *is at most* $\varepsilon^{-d_s^*}$.

Proof. — Indeed, the number of simple descendants P with $\max_I |P| > \varepsilon \max_I |P^*|$ is of order $\leqslant \varepsilon^{-d_s^0 + C\varepsilon_0}$, and the number of non-simple ones with $\max_I |P| > \varepsilon \max_I |P^*|$ is $\ll \varepsilon^{-d_s^*}$ from the proposition. □

8.4. *Transfer to parameter space.*

8.4.1. Our goal in this subsection will be to prove the following result, which expresses a transfer of the dimension estimate of Section 8.2 to parameter space.

Proposition **28.** — *Let* $\widetilde{I}$ *be a regular parameter interval. Let* (P^*, Q^*, n^*) *be an element of* $\mathcal{R}(\widetilde{I})$ *such that* Q^* *is* $\widetilde{I}$*-critical and*

$$(8.54) \qquad |Q^*| \leqslant \frac{1}{2}|\widetilde{I}|^{(1+\tau)(1-\eta)^{-1}}$$

for all $t \in \widetilde{I}$*. Then, the number of candidates* $I \subset \widetilde{I}$ *of the next level, such that* Q^* *is* I*-critical, is at most* $|\widetilde{I}|^{-\tau d_s^+}$*, 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 $|\widetilde{I}|^{-\tau}$. Proposition 28 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 28.

8.4.2. We make some general observations, that could have been made much earlier, but are only useful now.

Let (P, Q, n), (P_0, Q_0, n_0), (P'_0, Q'_0, n'_0), be elements of $\mathcal{R}(I)$ such that $P \subset P_s$, $Q'_0 \subset Q_0 \subset Q_u$, and $P'_0 \neq P_0$. From property (MP7) of the Markov partition (Section 3.4), it is easy to see that the condition (3.50) of Section 3.6.4 is satisfied. As explained in this subsection, we have then, for any $t \in I$

$$(8.55) \qquad \delta(Q_0, P) + C^{-1}|Q_0| \leqslant \delta(Q'_0, Ps),$$

$$(8.56) \qquad \delta_R(Q_0, P) + C^{-1}|Q_0| \leqslant \delta_R(Q'_0, P),$$

$$(8.57) \qquad \delta_L(Q_0, P) - C^{-1}|Q_0| \geqslant \delta_L(Q'_0, P),$$

$$(8.58) \qquad \delta_{LR}(Q_0, P) - C^{-1}|Q_0| \geqslant \delta_{LR}(Q'_0, P).$$

Let now (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 $P_0 \cap P_1 = \emptyset$. From the discussion in Section 3.6.2, we have either

$$(8.59) \qquad \begin{aligned} \delta_R(Q, P_0) &< \delta(Q, P_1), \\ \delta_{LR}(Q, P_0) &< \delta_L(Q, P_1), \end{aligned}$$

or the same inequalities after exchanging P_0 and P_1.

Proposition **29.** — *Assume that* (8.59) *holds (see Figure 8).*

1. *If* Q *and* P_1 *are* I*-separated, then* Q *and* P_0 *are* I*-separated.*
2. *If* Q *and* P_0 *are* I*-transverse, and* $|P_1|^{1-\eta} \leqslant |I|$ *for some* $t \in I$*, then* Q *and* P_1 *are* I*-transverse.*

Proof. — 1. Fix $t \in I$. We will assume that $G(Q \cap L_u) \cap P_1 \cap \Lambda$ is empty and show that $G(Q \cap L_u) \cap P_0 \cap \Lambda$ is also empty.

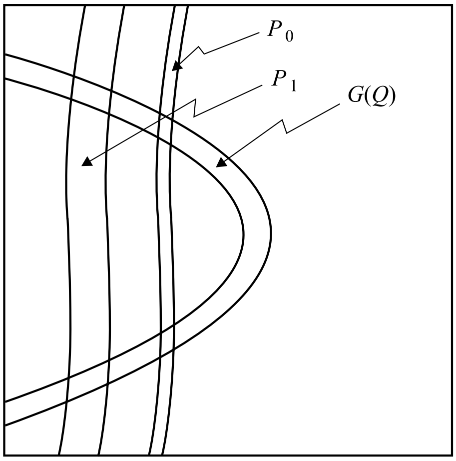

FIG. 8. — Proposition 29, part 2

Let γ a vertical-like curve which is the intersection of a decreasing sequence of simple descendants of P_1. The curve γ is contained in $W^s(\Lambda, \widehat{R}) \cap P_1$.

If γ does not intersect $G(Q \cap L_u)$, P_0 does also not intersect $G(Q \cap L_u)$ and the conclusion holds.

On the other hand, if the intersection of γ with $G(Q \cap L_u)$ is not connected, γ intersects the image $G(\gamma' \cap L_u)$, where γ' is any intersection of a decreasing sequence of simple descendants of Q; the intersection would be contained in $G(Q \cap L_u) \cap P_1 \cap \Lambda$, which is empty by hypothesis.

Therefore we can assume that the intersection γ_0 of γ with $G(Q \cap L_u)$ is non empty and connected. Let O_0 be the Jordan domain whose boundary is the union of γ_0 and the arc in the boundary of $G(Q \cap L_u)$ with the same endpoints than γ_0.

The intersection of P_0 with $G(Q \cap L_u)$ is contained in O_0. We will show that $O_0 \cap \Lambda$ is empty, which implies that $G(Q \cap L_u) \cap P_0 \cap \Lambda$ is empty.

Consider $(G \circ g^n)^{-1}(O_0) \subset P$. Part of the boundary of this Jordan domain is an horizontal segment contained in the boundary of P. The other part of the boundary is $(G \circ g^n)^{-1}(\gamma_0)$ which does not intersect $W^u(\Lambda, \widehat{R})$. In particular, it does not intersect L_s (in case $L_s \cap P$ is not empty), and it does not cross any $Q_{a,a'}$, $(a, a') \in \mathcal{B}$. Therefore, either it does not intersect any $Q_{a,a'}$ at all, in which case we can already conclude that $(G \circ g^n)^{-1}(O_0) \cap \Lambda$ is empty, or it intersects a single $Q_{a,a'}$. In this last case, there exists a simple child $Q^{(1)}$ of Q (defined by $(P_{a,a'}, Q_{a,a'}, 1) * (P, Q, n) = (P^{(1)}, Q^{(1)}, n+1)$) with the following property: let γ_1 be the subarc of γ_0 defined by $\gamma_1 = \gamma_0 \cap G(Q^{(1)} \cap L_u)$, and O_1 be the Jordan domain whose boundary is the union of γ_1 and the arc in the boundary of $G(Q^{(1)} \cap L_u)$ with the same endpoints than γ_1; then $O_0 \cap \Lambda = O_1 \cap \Lambda$.

We now apply to $Q^{(1)}, O_1, \gamma_1$ the same arguments that we used for Q, O_0, γ_0. Either we conclude that $O_1 \cap \Lambda$ is empty or we find a simple child $Q^{(2)}$ of $Q^{(1)}$, a subarc $\gamma_2 \subset \gamma_1$ and a Jordan domain O_2 whose boundary contains γ_2 such that $O_2 \cap \Lambda = O_1 \cap \Lambda$.

Iterating, either we conclude at some stage that $O_l \cap \Lambda$ is empty (and then we have $O_0 \cap \Lambda = O_l \cap \Lambda = \emptyset$), or we construct a decreasing sequence $Q^{(l)}$ of simple descendants of Q such that $\gamma \cap G(Q^{(l)} \cap L_u)$ is not empty for all l. But this is impossible, because the intersection γ' of the $Q^{(l)}$ is contained in $W^u(\Lambda, \widehat{R})$ and the intersection $\gamma \cap G(\gamma' \cap L_u)$ would be contained in $G(Q \cap L_u) \cap P_1 \cap \Lambda$.

2. Assume now that Q and P_0 are I-transverse, and that $|P_1|^{1-\eta} \leqslant |I|$ for some $t \in I$.

Let $\widetilde{I} \supset I$ and $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$, $(\widetilde{P}_0, \widetilde{Q}_0, \widetilde{n}_0) \in \mathcal{R}(\widetilde{I})$ be such that $\widetilde{Q} \supset Q$, $\widetilde{P}_0 \supset P_0$ and $\widetilde{Q} \,\overline{\pitchfork}_{\widetilde{I}}\, \widetilde{P}_0$ holds. If $P_1 \subset \widetilde{P}_0$, we immediately conclude that Q and P_1 are I-transverse. We assume, therefore, that $P_1 \cap \widetilde{P}_0 = \emptyset$; replacing (P_0, Q_0, n_0) by $(\widetilde{P}_0, \widetilde{Q}_0, \widetilde{n}_0)$, and (P, Q, n) by $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$, we can also assume that (P, Q, n), $(P_0, Q_0, n_0) \in \mathcal{R}(\widetilde{I})$ and $Q \,\overline{\pitchfork}_{\widetilde{I}}\, P_0$ holds.

Let $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1)$, be the element of $\mathcal{R}(\widetilde{I})$ with $P_1 \subset \widetilde{P}_1$ and smallest $\widetilde{P}_1$.

We will prove that Q and $\widetilde{P}_1$ are $\widetilde{I}$-transverse. We assume by contradiction that it is not the case. We have, for all $t \in \widetilde{I}$, that

$$\delta_{LR}(Q, \widetilde{P}_1) > \delta_{LR}(Q, P_0) \geqslant 2|\widetilde{I}|, \tag{8.60}$$

and also for some $t_0 \in \widetilde{I}$,

$$\delta_R(Q, \widetilde{P}_1) > \delta_R(Q, P_0) \geqslant 2|Q|^{1-\eta}. \tag{8.61}$$

Therefore, we must have, for all $t \in \widetilde{I}$, that

$$\delta_L(Q, \widetilde{P}_1) < 2|\widetilde{P}_1|^{1-\eta}. \tag{8.62}$$

We cannot have in this case $\widetilde{P}_1 = P_1$, because, for all $t \in I$,

$$\delta_L(Q, P_1) > \delta_{LR}(Q, P_0) \geqslant 2|\widetilde{I}|, \tag{8.63}$$

and (8.62), (8.63) together would contradict the hypothesis of the proposition. Therefore, $\widetilde{P}_1$ strictly contains P_1 and $\widetilde{I}$ strictly contains I. But, then, applying the structure theorem of Section 6.5 to the child of $\widetilde{P}_1$ which contains P_1, we obtain that $\widetilde{Q}_1$ is $\widetilde{I}$-critical. As $\widetilde{I}$ is β-regular, it then follows from (8.62), (8.63) and (3.39) that $\widetilde{P}_1$ is $\widetilde{I}$-transverse. This implies that there exists $(P', Q', n') \in \mathcal{R}(\widetilde{I})$ with $Q \cap Q' \neq \emptyset$ such that $Q' \pitchfork_{\widetilde{I}} \widetilde{P}_1$ holds.

When $Q \subset Q'$, it follows that $Q \pitchfork_{\widetilde{I}} \widetilde{P}_1$ holds.

When $Q' \subset Q$ and $P_0 \subset \widetilde{P}_1$, it follows by concavity from $Q' \pitchfork_{\widetilde{I}} \widetilde{P}_1$ and $Q \pitchfork_{\widetilde{I}} P_0$ that $Q \pitchfork_{\widetilde{I}} \widetilde{P}_1$ holds.

When $Q' \subset Q$ and $P_0 \cap \widetilde{P}_1 = \emptyset$, we consider a thin simple descendant P_1^* of $\widetilde{P}_1$ such that $|P_1^*|^{1-\eta} < |\widetilde{I}|$ for some $t_1 \in \widetilde{I}$. We have, as in (8.60), $\delta_{LR}(Q, P_1^*) \geqslant 2|\widetilde{I}|$ for all $t \in \widetilde{I}$, and, as in (8.61), $\delta_R(Q, P_1^*) \geqslant 2|Q|^{1-\eta}$ for some $t_0 \in \widetilde{I}$. But we have also, at $t_1 \in \widetilde{I}$, $\delta_L(Q, P_1^*) > \delta_{LR}(Q, P_0) \geqslant 2|\widetilde{I}| > 2|P_1^*|^{1-\eta}$. Therefore, $Q \pitchfork_{\widetilde{I}} P_1^*$ holds. We conclude by concavity from $Q' \pitchfork_{\widetilde{I}} \widetilde{P}_1$ and $Q \pitchfork_{\widetilde{I}} P_1^*$ that $Q \pitchfork_{\widetilde{I}} \widetilde{P}_1$ holds.

We thus obtain a contradiction in all cases. This proves that $Q \pitchfork_{\widetilde{I}} \widetilde{P}_1$ holds and thus also $Q \pitchfork_I P_1$. □

8.4.3. We now switch back to the setting of Proposition 28.

Let $(P, Q, n) \in \mathcal{R}(\widetilde{I})$ with $P \subset P_s$. We say that P is *eventually* $\widetilde{I}$*-transverse* to Q^* if there exists an $\widetilde{I}$-decomposition $(P_\alpha, Q_\alpha, n_\alpha)$ of P such that Q^* and P_α are $\widetilde{I}$-transverse for every α. We say that P is *eventually* $\widetilde{I}$*-*Q^**-critical* if it is neither $\widetilde{I}$-separated from Q^* nor eventually $\widetilde{I}$-transverse to Q^*.

Lemma **5.** — *If* P *is eventually* $\widetilde{I}$*-transverse to* Q^* *and* $2|P|^{1-\eta} \leqslant |\widetilde{I}|$ *holds for some* $t \in \widetilde{I}$, *then* $Q^* \pitchfork_{\widetilde{I}} P$ *holds.*

Proof. — This is an immediate consequence of Proposition 10 in Section 6.4. □

Lemma **6.** — *If* P *is eventually* $\widetilde{I}$*-*Q^**-critical, then* P *is* $\widetilde{I}$*-critical.*

Proof. — Assume on the contrary that P is $\widetilde{I}$-transverse. Let $(P_\alpha, Q_\alpha, n_\alpha)$ be an $\widetilde{I}$-decomposition of Q_u such that, for each α, Q_α and P are either $\widetilde{I}$-separated or $\widetilde{I}$-transverse.

If $Q^* \subset Q_\alpha$ for some α, Q^* and P would be $\widetilde{I}$-separated if Q_α and P are $\widetilde{I}$-separated, and $\widetilde{I}$-transverse if Q_α and P are $\widetilde{I}$-transverse.

If there exists α such that $Q_\alpha \subset Q^*$ and $Q_\alpha \pitchfork_{\widetilde{I}} P$ holds, then $Q^* \pitchfork_{\widetilde{I}} P$ also holds by Proposition 10 in Section 6.4.

In the remaining case, the $Q_\alpha \subset Q^*$ form an $\widetilde{I}$-decomposition of Q^* and they are all $\widetilde{I}$-separated from P; this imply that Q^* itself is $\widetilde{I}$-separated from P.

In all cases, we get a contradiction. The proof of the lemma is complete. □

Lemma **7.** — *If* P *is eventually* $\widetilde{I}$*-*Q^**-critical and* $|P| > |\widetilde{I}|^\beta$ *holds for some* $t \in \widetilde{I}$, *then some child of* P *is also eventually* $\widetilde{I}$*-*Q^**-critical.*

Proof. — We assume by contradiction that none of the children is eventually $\widetilde{I}$-Q^*-critical.

By Lemma 6 and Corollary 3 (in Section 5.6.5), P is $\widetilde{I}$-critical and thus $\widetilde{I}$-decomposable. If all children of P were eventually $\widetilde{I}$-transverse to Q^*, we would put together the corresponding $\widetilde{I}$-decompositions and obtain that P is eventually $\widetilde{I}$-transverse to Q^*. If all children of P were $\widetilde{I}$-separated from Q^*, P would be $\widetilde{I}$-separated from Q^* Therefore some child of P is eventually $\widetilde{I}$-transverse to Q^*, and some other child of P is $\widetilde{I}$-separated from Q^*.

We will show that Q^* is $\widetilde{I}$-transverse. We will construct an $\widetilde{I}$-decomposition $(P_\alpha, Q_\alpha, n_\alpha)$ of P_s such that every P_α is either $\widetilde{I}$-separated from Q^* or $\widetilde{I}$-transverse to Q^*.

Actually, it is sufficient to have an $\widetilde{I}$-decomposition such that every P_α is either $\widetilde{I}$-separated from Q^* or eventually $\widetilde{I}$-transverse to Q^*.

Starting from the trivial decomposition of P_s, we have at step i an $\widetilde{I}$-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 $\widetilde{I}$-critical and therefore $\widetilde{I}$-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.

The $P_\alpha^{(i)}$ which are children of P are either eventually $\widetilde{I}$-transverse to Q^* or $\widetilde{I}$-separated from Q^* by assumption.

The $P_\alpha^{(i)}$ which are disjoint from P may sit on one or the other side of P. On one side, we apply the first part of Proposition 29 to Q^*, $P_\alpha^{(i)}$ and a child of P which is $\widetilde{I}$-separated from Q^* to conclude that $P_\alpha^{(i)}$ is $\widetilde{I}$-separated from Q^*.

We claim that those on the other side are eventually $\widetilde{I}$-transverse to Q^*. Indeed, let $P_\alpha^{(i)}$ be such a rectangle. If it is $\widetilde{I}$-transverse to Q^*, we are done. Assume this is not the case. Let $\overline{P}_\alpha^{(i)}$ be a simple descendant of $P_\alpha^{(i)}$ such that $|\overline{P}_\alpha^{(i)}|^{1-\eta} < |\widetilde{I}|$ for some $t \in \widetilde{I}$. By Proposition 29 (part 2) applied to $\widetilde{I}$, Q^*, P_0, $\overline{P}_\alpha^{(i)}$, we have $Q^* \pitchfork_{\widetilde{I}} \overline{P}_\alpha^{(i)}$. As $P_\alpha^{(i)}$ is not $\widetilde{I}$-transverse to Q^*, it must be $\widetilde{I}$-critical (by Proposition 20 in Section 8.1), hence $\widetilde{I}$-decomposable. We replace $P_\alpha^{(i)}$ by its children and repeat the argument till the rectangles are thin enough to apply directly Proposition 29 (part 2) to $\widetilde{I}$, Q^*, P_0, $P_\alpha^{(i)}$.

This proves our claim and the proof of the lemma is complete. □

Lemma **8.** — *If* P_0, P_1 *are eventually* $\widetilde{I}$-Q^**-critical and disjoint, then we have* $|P_0| \leqslant C|\widetilde{I}|$, $|P_1| \leqslant C|\widetilde{I}|$ *for all* $t \in \widetilde{I}$.

Proof. — Exchanging P_0, P_1 if necessary, we can assume that (8.59) holds for Q^*, P_0, P_1.

From Lemma 7, we can find $(\widehat{P}_0, \widehat{Q}_0, \widehat{n}_0)$, $(\widehat{P}_1, \widehat{Q}_1, \widehat{n}_1)$ in $\mathcal{R}(\widetilde{I})$ with $\widehat{P}_0 \subset P_0$, $\widehat{P}_1 \subset P_1$, such that both $\widehat{P}_0$, $\widehat{P}_1$ are eventually $\widetilde{I}$-Q^*-critical and we have

$$(\mathbf{8.64}) \qquad |\widehat{P}_0| < |\widetilde{I}|^\beta, \qquad |\widehat{P}_1| < |\widetilde{I}|^\beta \quad \text{for all } t \in \widetilde{I}.$$

As $\widehat{P}_0$ is not $\widetilde{I}$-separated from Q^*, we must have

$$(\mathbf{8.65}) \qquad \delta_{LR}(Q^*, \widehat{P}_0) \geqslant 0$$

for some $t_0 \in \widetilde{I}$.

From (8.54) and (8.64), as $Q^* \overline{\pitchfork}_{\widetilde{I}} \widehat{P}_1$ does not hold, we must have

$$(\mathbf{8.66}) \qquad \delta_{LR}(Q^*, \widehat{P}_1) < 2|\widetilde{I}|$$

for some $t_1 \in I$.

Observe that Q^* is $\widetilde{I}$-critical by assumption and that P_0, P_1 are $\widetilde{I}$-critical by Lemma 6. From Corollary 8 in Section 7.6, it follows that, for $i = 0, 1$

$$(\mathbf{8.67}) \qquad \max_{\widetilde{I}} \delta_{LR}(Q^*, P_i) - \min_{\widetilde{I}} \delta_{LR}(Q^*, P_i) \leqslant 2|\widetilde{I}|,$$

$$(\mathbf{8.68}) \qquad \max_{\widetilde{I}} \delta_{L}(Q^*, P_i) - \min_{\widetilde{I}} \delta_{L}(Q^*, P_i) \leqslant 2|\widetilde{I}|.$$

From Proposition 18 in Section 7.7, for $i = 0, 1$, we have either $|P_i| < |\widetilde{I}|$ for all $t \in \widetilde{I}$ (the required conclusion), or

$$\max_{\widetilde{I}} |P_i| \leqslant \min_{\widetilde{I}} |P_i|. \tag{8.69}$$

Recall that, by (8.55)–(8.58), we have, for $i = 0, 1$, if $P_i \neq \widehat{P}_i$

$$\delta_{LR}(Q^*, P_i) > \delta_{LR}(Q^*, \widehat{P}_i) + C^{-1}|P_i|, \tag{8.70}$$

$$\delta_{L}(Q^*, P_i) < \delta_{L}(Q^*, \widehat{P}_i) - C^{-1}|P_i| \tag{8.71}$$

for any $t \in \widetilde{I}$.

If $P_0 = \widehat{P}_0$, we have $|P_0| < |\widetilde{I}|$ by (8.64). If $P_0 \neq \widehat{P}_0$, we have

$$\begin{aligned}
C^{-1}|P_0| &\leqslant C^{-1}|P_0| + \delta_{LR}(Q^*, \widehat{P}_0) \quad \text{at } t_0, \text{ from (8.63)} \\
&\leqslant \delta_{LR}(Q^*, P_0) \quad \text{at } t_0, \text{ from (8.68)} \\
&\leqslant \delta_{L}(Q^*, P_1) \quad \text{at } t_0 \\
&\leqslant \delta_{L}(Q^*, P_1) + C|\widetilde{I}| \quad \text{at } t_1, \text{ from (8.65)} \\
&\leqslant \delta_{LR}(Q^*, \widehat{P}_1) + C|\widetilde{I}| \quad \text{at } t_1 \\
&\leqslant C|\widetilde{I}| \quad \text{from (8.64)}.
\end{aligned}$$

If $P_1 = \widehat{P}_1$, we have $|P_1| < |\widetilde{I}|$ by (8.64). If $P_1 \neq \widehat{P}_1$, we have

$$\begin{aligned}
2|\widetilde{I}| - C^{-1}|P_1| &\geqslant \delta_{LR}(Q^*, \widehat{P}_1) - C^{-1}|P_1| \quad \text{at } t_1, \text{ from (8.64)} \\
&\geqslant \delta_{L}(Q^*, \widehat{P}_1) - C^{-1}|P_1| \quad \text{at } t_1 \\
&\geqslant \delta_{L}(Q^*, P_1) \quad \text{at } t_1, \text{ from (8.69)} \\
&\geqslant \delta_{L}(Q^*, P_1) - C|\widetilde{I}| \quad \text{at } t_0, \text{ from (8.65)} \\
&\geqslant \delta_{LR}(Q^*, \widehat{P}_0) - C|\widetilde{I}| \quad \text{at } t_0 \\
&\geqslant -C|\widetilde{I}| \quad \text{from (8.63)}.
\end{aligned}$$

We have proven the required estimates. □

8.4.4. Consider the set Π of elements $(P, Q, n) \in \mathcal{R}(\widetilde{I})$ which are eventually $\widetilde{I}$-Q^*-critical, satisfy

$$|P| \leqslant |\widetilde{I}|^{1+\tau} \tag{8.72}$$

for all $t \in \widetilde{I}$ and are maximal (in P) with respect to these two properties.

Lemma **9.** — *We have*

$$\#\Pi \leqslant |\widetilde{I}|^{-\tau d_s^+},$$

where $d_s^+ = d_s^0 + C\eta\tau^{-1}$ *is as in the statement of Proposition* 28.

Proof. — Assume that Π is non-empty. From Lemma 8, there exists $C_0 > 0$ and a unique element $(P_0, Q_0, n_0) \in \mathcal{R}(\widetilde{I})$ with the following properties:

- $P \subset P_0$ for all $(P, Q, n) \in \Pi$
- $|P_0| > C_0|\widetilde{I}|$ for some $t \in \widetilde{I}$
- every child P_1 of P_0 which contains a rectangle P with $(P, Q, n) \in \Pi$ satisfies $|P_1| \leqslant C_0|\widetilde{I}|$ for all $t \in \widetilde{I}$.

As P_0 is eventually $\widetilde{I}$-Q^*-critical, P_0 is $\widetilde{I}$-critical by Lemma 6.

There are two kind of elements $(P, Q, n) \in \Pi$:

- those such that P is a child of P_0; the number of such elements is at most $|\widetilde{I}|^{-c\eta}$ by Proposition 25;
- those such that the parent $\widetilde{P}$ of P is contained in some child P_1 of P_0.

In this last case, from the definition of Π we have

$$(\mathbf{8.73}) \qquad |\widetilde{P}| > |\widetilde{I}|^{1+\tau} \quad \text{for some } t \in \widetilde{I}.$$

As P is $\widetilde{I}$-critical by Lemma 6, $\widetilde{P}$ is also $\widetilde{I}$-critical, hence $\widetilde{I}$-special and we have, from Proposition 18 in Section 7.7, that

$$(\mathbf{8.74}) \qquad |\widetilde{P}| > C^{-1}|\widetilde{I}|^{1+\tau} \quad \text{for all } t \in \widetilde{I}.$$

Let P_1 be a child of P_0; we have $|P_1| \leqslant C_0|\widetilde{I}|$ for all $t \in \widetilde{I}$ by the definition of P_0. For given P_1, the number of possible $\widetilde{P}$ is therefore bounded by $C|\widetilde{I}|^{-\tau d_s^*}$ from Corollary 13, with $d_s^* = d_s^0 + \varepsilon_0^{\frac{1}{5}d_s^0}$.

The number of P for given $\widetilde{P}$, and the number of children P_1 of P_0, are both bounded from Proposition 25 by $|\widetilde{I}|^{-c\eta}$. Summing up, the total number of elements of Π is bounded by

$$|\widetilde{I}|^{-c\eta} + |\widetilde{I}|^{-2c\eta}(C|\widetilde{I}|^{\tau})^{-d_s^*},$$

in accordance with the statement of Lemma 9, choosing appropriately the constant in the definition of d_s^+. □

8.4.5. *Proof of Proposition 28.* — By Lemma 6 and Corollary 3 (in Section 5.6.5), if $(P, Q, n) \in \mathcal{R}(\tilde{I})$ is such that P is eventually $\tilde{I}$-Q^*-critical, and $|P| \geqslant |\tilde{I}|^\beta$ for some $t \in \tilde{I}$, then P 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. This proves the claim.

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 is eventually I-Q^*-critical. Together with Lemma 9, this will imply the statement of Proposition 28 (after modifying the value of the constant in the definition of d_s^+).

Both P (by Lemma 6) and Q^* (by assumption) are $\tilde{I}$-critical, hence $\tilde{I}$-special. By Corollary 8 in Section 7.6, we have, for all $t \in \tilde{I}$

$$\left|\frac{d}{dt}\delta_{LR}(Q^*, P) - 1\right| \leqslant C\varepsilon_0^{\frac{1}{2}}. \quad (\mathbf{8.75})$$

If, for all $t \in I$, we have

$$\delta_{LR}(Q^*, P) < 0, \quad (\mathbf{8.76})$$

then P is I-separated from Q^*.

We claim that if, for all $t \in I$, we have

$$\delta(Q^*, P) > 2|I|, \quad (\mathbf{8.77})$$

then P is eventually I-transverse to Q^*. As we have $|P| \leqslant |I|$ from the definition of Π and $|Q| < |I|$ from the assumption of the proposition, we have, for all $t \in \tilde{I}$

$$\delta_{LR}(Q^*, P) < \delta(Q^*, P) + C(|P| + |Q^*|) < \delta(Q^*, P) + C|I|, \quad (\mathbf{8.78})$$

which allows to conclude.

Finally, we prove the claim.

Observe first that, if $(P', Q', n') \in \mathcal{R}(I)$ satisfies $P' \subset P$, $|P'| \leqslant \frac{1}{2}|\tilde{I}|^{(1+\tau)(1-\eta)^{-1}}$ for all $t \in \tilde{I}$, then $Q^* \overline{\pitchfork}_I P'$ holds. Indeed, $\delta_{LR}(Q^*, P')$, $\delta_L(Q^*, P')$, $\delta_R(Q^*, P')$ are all larger than $\delta(Q^*, P)$; then (T3) follows from (8.75) above; as both $|Q^*|^{1-\eta}$, $|P|^{1-\eta}$ are smaller than $|I|$, (T1) and (T2) also follow from (8.75).

On the other hand, if $(P', Q', n') \in \mathcal{R}(I)$ satisfies $P' \subset P$ and P' is not I-transverse to Q^*, from the observation (applied to a thin simple descendant of P') and Proposition 20, we deduce that P' is I-critical hence I-decomposable, and we can replace it by its children. Starting with P, we iterate the process till getting rectangles thin enough for the observation to apply. This proves the claim and thus also Proposition 28.

9. Strong regularity and parameter selection

As it was mentioned in Section 4.5, regularity is a rather qualitative property which is not appropriate for the quantitative estimates needed for parameter selection. We will introduce in this section, a stronger quantitative property, that we call *strong regularity*.

Think of each bicritical element as a return from the "critical region" to itself. We want to control the number of bicritical elements of a given width (including, of course, that there are no "fat" bicritical elements). In order to do this, we also need a control on the size of the critical region itself.

In the whole of Section 9, we fix a parameter interval $\widetilde{I}$ which is assumed by induction to be strongly regular (the definition is given at the end of Section 9.2). In Section 9.3, we check that strong regularity implies β-regularity for an appropriate $\beta > 1$. Therefore the properties proven in Sections 5–8 will be applicable. The aim of Section 9 is to estimate how many candidates $I \subset \widetilde{I}$ fail to be strongly regular.

In Section 9.1, the estimates (SR1), (SR2) on the size of the critical region are presented. In Section 9.2, the estimates (SR3) on the number of bicritical elements are introduced. This leads to the definition of strong regularity at the end of this Subsection, namely that the full set of estimates (SR1), (SR2), (SR3) must be satisfied, and to the choice of β in Section 9.3, where condition (H4) of Section 1.2 is finally explained. It is also proven in Section 9.3 that strong regularity implies β-regularity. In Section 9.4, we check that the starting interval I_0 is strongly regular. Then, in Section 9.5, we estimate how many candidates $I \subset \widetilde{I}$ fail to satisfy (SR1). Condition (SR2) is more delicate and is dealt with in Section 9.6.

In the rest of Sections 9, 9.7–9.13, we estimate the number of candidates which fail to satisfy (SR3); a general overview of the strategy to do this is first presented in Section 9.7. The conclusions of the process of parameter selection are presented in Section 9.14.

9.1. *Partitions and size of the critical locus.*

Definition **3.** — *We denote by* $\mathcal{C}_+(I)$ *the set of* (P, Q, n) *in* $\mathcal{R}(I)$ *such that* P *is* I-*critical,* $|P| \leqslant |I|^{1+\tau}$ *for all* $t \in I$, *and* P *is maximal with this property: the parent* $\widetilde{P}$ *of* P *satisfies* $|\widetilde{P}| > |I|^{1+\tau}$ *for some* $t \in I$.

Obviously, if (P, Q, n), (P', Q', n') are distinct elements in $\mathcal{C}_+(I)$, P and P' are disjoint. Moreover, if $(\widehat{P}, \widehat{Q}, \widehat{n})$ belongs to $\mathcal{R}(I)$, $\widehat{P}$ is I-critical, and $|\widehat{P}| \leqslant |I|^{1+\tau}$ for all $t \in I$, there is a unique $(P, Q, n) \in \mathcal{C}_+(I)$ such that $\widehat{P} \subset P$.

Exchanging P's and Q's, we define $\mathcal{C}_-(\mathrm{I})$ in a similar way. The sets $\mathcal{C}_+(\mathrm{I})$, $\mathcal{C}_-(\mathrm{I})$ correspond to the I-critical locus at the $|\mathrm{I}|^{1+\tau}$ scale.

We will need in the sequel to consider $\widehat{\mathrm{I}}$-criticality (for some parameter interval $\widehat{\mathrm{I}} \supset \mathrm{I}$) for rectangles in $\mathcal{R}(\mathrm{I})$ but not in $\mathcal{R}(\widehat{\mathrm{I}})$.

Definition **4.** — *Let* $\widehat{\mathrm{I}} \supset \mathrm{I}$ *be parameter intervals, and let* $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}(\mathrm{I})$. *We say that* P *is thin* $\widehat{\mathrm{I}}$*-critical if there exists* $(\widehat{\mathrm{P}}, \widehat{\mathrm{Q}}, \widehat{n}) \in \mathcal{R}(\widehat{\mathrm{I}})$ *with* $\mathrm{P} \subset \widehat{\mathrm{P}}$, $\widehat{\mathrm{P}}$ *is* $\widehat{\mathrm{I}}$*-critical and*

$$\min_{\widehat{\mathrm{I}}} |\widehat{\mathrm{P}}|^{1-\eta} \leqslant 2|\widehat{\mathrm{I}}|.$$

The notion is useful in connection in Proposition 22 of Section 8.1, as in the following

Lemma **10.** — *Let* I *be a candidate parameter interval (with regular parent* $\widetilde{\mathrm{I}}$*), and let* $(\mathrm{P}, \mathrm{Q}, n)$ *be an element of* $\mathcal{R}(\mathrm{I})$ *which is not the restriction of an element of* $\mathcal{R}(\widetilde{\mathrm{I}})$. *Let* $k > 0$ *and* $(\mathrm{P}_i, \mathrm{Q}_i, n_i)$, *for* $0 \leqslant i \leqslant k$, *be the elements of* $\mathcal{R}(\widetilde{\mathrm{I}})$ *given by the structure theorem (of Section 6.5). Then* Q_i *and* P_{i+1} *are thin* $\widetilde{\mathrm{I}}$*-critical for* $0 \leqslant i < k$.

Proof. — As Q_i is I-transverse but not $\widetilde{\mathrm{I}}$-transverse to P_{i+1}, this follows from Proposition 22 in Section 8.1. □

Definition **5.** — *Let* $\widetilde{\mathrm{I}}$ *be a regular interval,* $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{C}_+(\widetilde{\mathrm{I}})$. *We denote by* $\mathrm{C}r(\mathrm{P})$ *the set of candidates* $\mathrm{I} \subset \widetilde{\mathrm{I}}$ *such that* P *contains a thin* I*-critical rectangle. We define symmetrically* $\mathrm{C}r(\mathrm{Q})$ *for* $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{C}_-(\widetilde{\mathrm{I}})$.

Lemma **11.** — *Let* $\widehat{\mathrm{I}}$ *be the largest parameter interval containing* $\widetilde{\mathrm{I}}$ *with* $|\widehat{\mathrm{I}}|^{\beta} \leqslant |\widetilde{\mathrm{I}}|^{(1+\tau)^2}$. *Let* $\mathrm{I} \subset \widetilde{\mathrm{I}}$ *be a candidate interval and let* $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{C}_+(\mathrm{I})$. *Then* $(\mathrm{P}, \mathrm{Q}, n)$ *belongs to* $\mathcal{R}(\widehat{\mathrm{I}})$, *and there exists* $(\widetilde{\mathrm{P}}, \widetilde{\mathrm{Q}}, \widetilde{n}) \in \mathcal{C}_+(\widetilde{\mathrm{I}})$ *with* $\mathrm{P} \subset \widetilde{\mathrm{P}}$ *and* $\mathrm{I} \in \mathrm{C}r(\widetilde{\mathrm{P}})$.

Proof. — We first prove that $(\mathrm{P}, \mathrm{Q}, n)$ belongs to $\mathcal{R}(\widehat{\mathrm{I}})$. Let $\mathrm{I}' \supset \mathrm{I}$ be a parameter interval distinct from I_0 with parent $\widetilde{\mathrm{I}}'$; assume that $(\mathrm{P}, \mathrm{Q}, n)$ belongs to $\mathcal{R}(\mathrm{I}')$ but not to $\mathcal{R}(\widetilde{\mathrm{I}}')$. By the structure theorem of Section 6.5, the element $(\mathrm{P}_0, \mathrm{Q}_0, n_0) \in \mathcal{R}(\widetilde{\mathrm{I}}')$ with P_0 smallest containing P is $\widetilde{\mathrm{I}}'$-bicritical and satisfy $|\mathrm{P}_0| < |\widetilde{\mathrm{I}}'|^{\beta}$ for all $t \in \widetilde{\mathrm{I}}'$. This contradicts the maximality of P with respect to the property $\max_{\mathrm{I}} |\mathrm{P}| \leqslant |\mathrm{I}|^{1+\tau}$ if $|\widetilde{\mathrm{I}}'|^{\beta} \leqslant |\mathrm{I}|^{1+\tau}$ and shows that $(\mathrm{P}, \mathrm{Q}, n)$ belongs to $\mathcal{R}(\widehat{\mathrm{I}})$.

In particular, $(\mathrm{P}, \mathrm{Q}, n)$ belongs to $\mathcal{R}(\widetilde{\mathrm{I}})$ and P is $\widetilde{\mathrm{I}}$-special. From Proposition 18 in Section 7.7, we have $|\mathrm{P}| < \mathrm{C}|\mathrm{I}|^{1+\tau} < |\widetilde{\mathrm{I}}|^{1+\tau}$ for all $t \in \widetilde{\mathrm{I}}$. Therefore, there exists $(\widetilde{\mathrm{P}}, \widetilde{\mathrm{Q}}, \widetilde{n}) \in \mathcal{C}_+(\widetilde{\mathrm{I}})$ with $\mathrm{P} \subset \widetilde{\mathrm{P}}$. Finally, we observe that P is thin I-critical and therefore I belongs to $\mathrm{C}r(\widetilde{\mathrm{P}})$. □

We will state several inequalities related to the size of the sets $\mathcal{C}_+(\mathrm{I})$, $\mathcal{C}_-(\mathrm{I})$. All these inequalities are part of the definition of strong regularity: they have to be satisfied by a strongly regular parameter interval. In Sections 9.5, 9.6, we will see that they are satisfied by most candidates in a strongly regular parameter interval $\widetilde{I}$. These estimates are used in Sections 9.12, 9.13 in the control of the number of bicritical elements.

We will control the cardinalities of $\mathcal{C}_+(I)$, $\mathcal{C}_-(I)$, through:

$$(SR1)_s \qquad \#\mathcal{C}_+(I) \leqslant C\left(\frac{|I|}{\varepsilon_0}\right)^{\sigma} \varepsilon_0^{-\tau d_s^0},$$

$$(SR1)_u \qquad \#\mathcal{C}_-(I) \leqslant C\left(\frac{|I|}{\varepsilon_0}\right)^{\sigma} \varepsilon_0^{-\tau d_u^0}.$$

The exponent σ will be completely specified in Section 9.5. It is very close to $1 - d_s^0 - d_u^0$ when $\tau, \eta, \varepsilon_0$ are small.

We need also to control $\mathcal{C}_+(I)$, $\mathcal{C}_-(I)$ in another, more complicated, way.

$$(SR2)_s \qquad \sum_{(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(I)} \max_I |Q_\alpha|^{\rho_u} \leqslant C|Q_s|^{\rho_u}\left(\frac{|I|}{\varepsilon_0}\right)^{\sigma_u},$$

$$(SR2)_u \qquad \sum_{(P_\omega, Q_\omega, n_\omega) \in \mathcal{C}_-(I)} \max_I |P_\omega|^{\rho_s} \leqslant C|P_u|^{\rho_s}\left(\frac{|I|}{\varepsilon_0}\right)^{\sigma_s}.$$

The exponents $\rho_s, \rho_u, \sigma_s, \sigma_u$ will be specified in Sections 9.6 and 9.14. When $\tau, \eta, \varepsilon_0$ are small, they are respectively close to $d_s^0, d_u^0, 1 - d_s^0, 1 - d_u^0$. We write $|P_u|$, $|Q_s|$ for the maximum of these quantities over I_0.

We actually need a stronger version of $(SR2)_s$, better suited for induction purposes. Let $I \subset \widetilde{I}$ be a candidate interval.

Definition **6.** — *We denote by* $\widehat{\mathcal{C}}_+(I)$ *to be the set of* $(P, Q, n) \in \mathcal{C}_+(\widetilde{I})$ *such that* $I \in Cr(P)$. *We define symmetrically* $\widehat{\mathcal{C}}_-(I)$. *We also define* $\widehat{\mathcal{C}}_+(I_0) := \{(P_s, Q_s, n_s)\}$, $\widehat{\mathcal{C}}_-(I_0) := \{(P_u, Q_u, n_u)\}$.

Let $((P_i, Q_i, n_i))_i$ *be a finite family of elements of* $\mathcal{R}(I)$ *with the* P_i *disjoint and each* P_i *contained in some* P *with* $(P, Q, n) \in \widehat{\mathcal{C}}_+(I)$. *We ask that, for any such family*

$$(SR2)'_s \qquad \sum_i \max_I |Q_i|^{\rho_u} \leqslant C|Q_s|^{\rho_u}\left(\frac{|I|}{\varepsilon_0}\right)^{\sigma_u}.$$

We define symmetrically $(SR2)'_u$.

Observe that, by Lemma 11, $(SR2)_s$ is a consequence of $(SR2)'_s$.

The heuristics behind $(SR2)_s$ is the following: in the mean, one expects that elements of $\mathcal{R}(I)$ more or less satisfy

$$(\mathbf{9.1}) \qquad |P|^{d_s^0} \sim |Q|^{d_u^0}$$

and, for $(P, Q, n) \in \widehat{\mathcal{C}}_+(I)$, one should have

$$(\mathbf{9.2}) \qquad |P| \sim |I|$$

which explains the relation between $(SR1)_s$ and $(SR2)_s$.

9.2. *Classes and number of bicritical rectangles.* — Once 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 **7.** — *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* $|P| \geqslant x$ *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* $|Q| \geqslant x$ *for some* $t \in I$.

At this point, we have to break the symmetry between past and future, P's and Q's, stable and unstable directions: 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 $d_s^0 \geqslant d_u^0$ (and $d_s^0 + d_u^0 \geqslant 1$). The case $d_u^0 \geqslant d_s^0$ is obviously symmetric.

For I, I_α, I_ω, x as above we want to have, for a fixed large enough constant A (we will see in Section 9.14 that $A = 3$ is a convenient choice)

$$(SR3)_s \qquad \#Bi_+(I, I_\alpha, I_\omega; x) \leqslant \varepsilon_0^{-A\tau} B,$$

with

$$B = \max(B_0, B_1), \tag{9.3}$$

$$B_0 = \Big(\frac{x}{\varepsilon_0 |P_u|}\Big)^{-\rho_0} \Big(\frac{|I_\alpha|}{\varepsilon_0}\Big)^{\sigma_0+\sigma_1} \Big(\frac{|I_\omega|}{\varepsilon_0}\Big)^{\sigma_0}, \tag{9.4}$$

$$B_1 = \Big(\frac{x}{\varepsilon_0 |P_u|}\Big)^{-\rho_1} \Big(\frac{|I_\alpha|}{\varepsilon_0}\Big)^{\sigma_1} \Big(\min\Big(\frac{|I_\alpha|}{\varepsilon_0}, \frac{|I_\omega|}{\varepsilon_0}\Big)\Big)^{\sigma_0}. \tag{9.5}$$

Here $|P_u|$ denotes the supremum over I_0 of the width of P_u: the exponents ρ_0, ρ_1, σ_0, σ_1 will be specified more precisely later, but anyway they satisfy

$$\rho_0 = d_s^0 + o(1), \tag{9.6}$$

$$\rho_1 = \frac{d_s^0}{d_s^0 + d_u^0}(2d_s^0 + d_u^0 - 1) + o(1), \tag{9.7}$$

$$\sigma_0 = 1 - d_s^0 + o(1), \tag{9.8}$$

$$\sigma_1 = d_s^0 - d_u^0 + o(1). \tag{9.9}$$

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 $\mathrm{B}i_-$ sets, we should have:

$$\#\mathrm{B}i_-(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x) \leqslant \varepsilon_0^{-\mathrm{A}\tau}\mathrm{B}', \tag*{(SR3)$_u$}$$

with

$$\mathrm{B}' = \max(\mathrm{B}'_0, \mathrm{B}'_1), \tag{9.10}$$

$$\mathrm{B}'_0 = \left(\frac{x}{\varepsilon_0|\mathrm{Q}_s|}\right)^{-\rho'_0}\left(\frac{|\mathrm{I}_\alpha|}{\varepsilon_0}\right)^{\sigma_0+\sigma_1}\left(\frac{|\mathrm{I}_\omega|}{\varepsilon_0}\right)^{\sigma_0}, \tag{9.11}$$

$$\mathrm{B}'_1 = \left(\frac{x}{\varepsilon_0|\mathrm{Q}_s|}\right)^{-\rho'_1}\left(\frac{|\mathrm{I}_\alpha|}{\varepsilon_0}\right)^{\sigma_1}\left(\min\left(\frac{|\mathrm{I}_\alpha|}{\varepsilon_0}, \frac{|\mathrm{I}_\omega|}{\varepsilon_0}\right)\right)^{\sigma_0}, \tag{9.12}$$

$$\rho'_0 = \frac{d_u^0}{d_s^0}\rho_0 = d_u^0 + o(1), \tag{9.13}$$

$$\rho'_1 = \frac{d_u^0}{d_s^0}\rho_1 = \frac{d_u^0}{d_s^0 + d_u^0}(2d_s^0 + d_u^0 - 1) + o(1). \tag{9.14}$$

Observe that the formulas (9.7), (9.14) for ρ_1, ρ'_1 are *not* symmetric.

Definition **8.** — *A parameter interval* I *is strongly regular if its parent is (when* $\mathrm{I} \neq \mathrm{I}_0$*) and if it satisfies the conditions* (SR1), (SR2)′ *(hence also* (SR2)*) of Section* 9.1 *and* (SR3)$_s$, (SR3)$_u$ *for all* $\mathrm{I}_\alpha \supset \mathrm{I}$, $\mathrm{I}_\omega \supset \mathrm{I}$, $0 < x < 1$.

Remark **10.**

1. At this point, the definition of strong regularity is not complete because the exponents ρ_0, ρ_1, ρ'_0, ρ'_1, σ_0, σ_1, A have not been completely specified. These exponents should be viewed for the present time as parameters constrained by (9.6)–(9.9) and (9.13), (9.14).
2. The inequalities (SR3)$_s$, (SR3)$_u$ form a family parametrized not only by I, but also by the parameter intervals $\mathrm{I}_\alpha \supset \mathrm{I}$ and $\mathrm{I}_\omega \supset \mathrm{I}$ and the real number $1 > x > 0$. Because each inequality, at least when $\mathrm{I} = \mathrm{I}_\alpha$ or $\mathrm{I} = \mathrm{I}_\omega$, is only obtained after parameter selection, we will discretize the continuous variable x by considering only the values $x = 2^{-l}$, $l \geqslant 0$. There is still an infinite number of inequalities, but we will see that they are trivially satisfied if l is large enough.

9.3. *Phase transition and the choice of β.* — We comment on the estimate $(\mathrm{SR3})_s$. First, observe that B does not depend on I. The reason will appear in Section 9.8 when we show that most elements in $Bi_+(I, I_\alpha, I_\omega; x)$ belong actually to $\mathcal{R}(\widetilde{I})$.

From the formulas (9.6), (9.7), we have

$$\rho_1 < \rho_0. \tag{9.15}$$

Set

$$x_{cr} := \varepsilon_0 |P_u| \Big(\max\Big(\frac{|I_\alpha|}{\varepsilon_0}, \frac{|I_\omega|}{\varepsilon_0} \Big) \Big)^{\frac{\sigma_0}{\rho_0 - \rho_1}}. \tag{9.16}$$

Then, we have $B = B_0$ for $x \leqslant x_{cr}$ and $B = B_1$ for $x \geqslant x_{cr}$: this is the "phase transition" mentioned earlier. Roughly speaking, the reason for this phase transition is that, when $I = I_\alpha = I_\omega$, we are able to eliminate, in the scale transition from $\widetilde{I}$ to I, more bicritical elements of small width ($x < x_{cr}$) than of large width ($x > x_{cr}$).

We have

$$\rho_0 - \rho_1 = \frac{d_s^0(1 - d_s^0)}{d_s^0 + d_u^0} + o(1), \tag{9.17}$$

$$\frac{\sigma_0}{\rho_0 - \rho_1} = \frac{d_s^0 + d_u^0}{d_s^0} + o(1) > 1. \tag{9.18}$$

For $x = x_{cr}$, we have

$$B = B_{cr} := \Big(\frac{|I_\alpha|}{\varepsilon_0} \Big)^{\sigma_0 + \sigma_1} \Big(\frac{|I_\omega|}{\varepsilon_0} \Big)^{\sigma_0} \Big(\max\Big(\frac{|I_\alpha|}{\varepsilon_0}, \frac{|I_\omega|}{\varepsilon_0} \Big) \Big)^{-\frac{\rho_0 \sigma_0}{\rho_0 - \rho_1}}. \tag{9.19}$$

Assume $I_\alpha = I_\omega$; we then have

$$B_{cr} = \Big(\frac{|I_\alpha|}{\varepsilon_0} \Big)^{\sigma_1 + \sigma_0 \frac{\rho_0 - 2\rho_1}{\rho_0 - \rho_1}}. \tag{9.20}$$

Here, the exponent satisfies

$$\sigma_1 + \sigma_0 \frac{\rho_0 - 2\rho_1}{\rho_0 - \rho_1} = 2 - 2d_s^0 - 2d_u^0 + o(1) < 0. \tag{9.21}$$

As $|I_\alpha| \leqslant \varepsilon_0$, we have $B_{cr} \geqslant 1$. As B is a decreasing function of x, we have $\varepsilon_0^{-A\tau} B < 1$ (in which case $(\mathrm{SR3})_s$ means that the Bi_+ set is empty!) iff $B_1 < \varepsilon_0^{A\tau}$ which corresponds to

$$x > \overline{x} := \varepsilon_0^{1 + A\tau/\rho_1} |P_u| \Big(\frac{|I_\alpha|}{\varepsilon_0} \Big)^{\frac{\sigma_0 + \sigma_1}{\rho_1}}. \tag{9.22}$$

The exponent here satisfies

$$(\mathbf{9.23})\qquad \frac{\sigma_0+\sigma_1}{\rho_1} = \frac{1-d_u^0}{d_s^0}\frac{d_s^0+d_u^0}{2d_s^0+d_u^0-1} + o(1).$$

We are finally able to justify the assumption (H4) of our Main Theorem stated in Section 1.2! Indeed, with $d_s^0 \geqslant d_u^0$, it means that

$$(\mathbf{H4})\qquad 2(d_s^0)^2 + (d_u^0)^2 + 2d_s^0 d_u^0 < 2d_s^0 + d_u^0$$

and this is exactly what is needed to guarantee that

$$(\mathbf{9.24})\qquad \frac{\sigma_0+\sigma_1}{\rho_1} > 1.$$

The discussion for $(\mathrm{SR3})_u$ is similar; the critical threshold is

$$(\mathbf{9.25})\qquad x'_{cr} := \varepsilon_0 |Q_s| \left(\max\left(\frac{|I_\alpha|}{\varepsilon_0}, \frac{|I_\omega|}{\varepsilon_0}\right)\right)^{\frac{\sigma_0}{\rho'_0-\rho'_1}},$$

with

$$(\mathbf{9.26})\qquad \rho'_0 - \rho'_1 = \frac{d_u^0(1-d_s^0)}{d_s^0+d_u^0} + o(1) = \frac{d_u^0}{d_s^0}(\rho_0-\rho_1),$$

$$(\mathbf{9.27})\qquad \frac{\sigma_0}{\rho'_0-\rho'_1} = \frac{\sigma_0}{\rho_0-\rho_1}\frac{d_s^0}{d_u^0} = \frac{d_s^0+d_u^0}{d_u^0} + o(1) > 1.$$

When $I_\alpha = I_\omega$, we have

$$(\mathbf{9.28})\qquad B'_{cr} := \left(\frac{|I_\alpha|}{\varepsilon_0}\right)^{\sigma_1+\sigma_0\frac{\rho'_0-2\rho'_1}{\rho'_0-\rho'_1}} = B_{cr} \geqslant 1.$$

Thus, we have $\varepsilon_0^{-A\tau} B' < 1$ iff

$$(\mathbf{9.29})\qquad x > \overline{x}' := \varepsilon_0^{1+A\tau/\rho'_1} |Q_s| \left(\frac{|I_\alpha|}{\varepsilon_0}\right)^{\frac{\sigma_0+\sigma_1}{\rho'_1}}.$$

We have here

$$(\mathbf{9.30})\qquad \frac{\sigma_0+\sigma_1}{\rho'_1} = \frac{\sigma_0+\sigma_1}{\rho_1}\frac{d_s^0}{d_u^0} \geqslant \frac{\sigma_0+\sigma_1}{\rho_1}.$$

Choice of β. — We will choose the constant β (related to the regularity property) in order to have

$$1 < \beta < \frac{1-d_u^0}{d_s^0}\frac{d_s^0+d_u^0}{2d_s^0+d_u^0-1} \tag{9.31}$$

and also, from Proposition 4 in Section 5.6.4

$$\beta < 1 + \min(\omega_s, \omega_u). \tag{9.32}$$

From (9.31), we will have

$$\beta < \frac{\sigma_0+\sigma_1}{\rho_1} \leqslant \frac{\sigma_0+\sigma_1}{\rho_1'}. \tag{9.33}$$

Then, in (9.22), (9.29), we will have

$$\bar{x} < |\mathrm{I}_\alpha|^\beta, \qquad \bar{x}' < |\mathrm{I}_\alpha|^\beta. \tag{9.34}$$

Proposition **30.** — *If a candidate interval satisfies* $(\mathrm{SR3})_s$ *and* $(\mathrm{SR3})_u$, *then it is β-regular. In particular, strong regularity implies β-regularity.*

Proof. — We argue by induction on the level of the parameter interval. For the starting interval I_0, we already know from Proposition 4 that it is β-regular (independently of $(\mathrm{SR3})_s$, $(\mathrm{SR3})_u$). Assume that $\mathrm{I} \neq \mathrm{I}_0$ satisfies $(\mathrm{SR3})_s$, $(\mathrm{SR3})_u$ and that $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}(\mathrm{I})$ is I-bicritical. Assume also, for instance, that

$$\max_{\mathrm{I}} |\mathrm{Q}| \leqslant \max_{\mathrm{I}} |\mathrm{P}| \tag{9.35}$$

and, by contradiction that

$$\max_{\mathrm{I}} |\mathrm{P}| \geqslant |\mathrm{I}|^\beta. \tag{9.36}$$

From Corollary 5 in Section 6.6, we know that $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}(\widetilde{\mathrm{I}})$ ($\widetilde{\mathrm{I}}$ being the parent of I). As $(\mathrm{P}, \mathrm{Q}, n)$ is $\widetilde{\mathrm{I}}$-bicritical, we must have, by the induction hypothesis

$$\max_{\widetilde{\mathrm{I}}} |\mathrm{P}| < |\widetilde{\mathrm{I}}|^\beta. \tag{9.37}$$

Therefore, P would be thin I-critical; similarly Q would be thin I-critical. But in view of (9.34), the estimate $(\mathrm{SR3})_s$ says that such a $(\mathrm{P}, \mathrm{Q}, n)$ satisfying (9.36) does not exist. □

9.4. *The starting interval.* — We establish in this subsection the starting point of an inductive construction of strongly regular parameters. We assume that the ρ exponents satisfy

$$\rho_u > d_u^0 + C\varepsilon_0, \qquad \rho_s > d_s^0 + C\varepsilon_0, \qquad \rho_0 > d_s^0 + C\varepsilon_0, \qquad \rho_0' > d_u^0 + C\varepsilon_0, \tag{9.38}$$

with C large enough. The exponents σ do not come into play for the starting interval.

Proposition **31.** — *The starting interval is strongly regular.*

Proof. — We start by checking (SR1) for I_0.

The set $\mathcal{C}_+(I_0)$ consists of the elements $(P, Q, n) \in \mathcal{R}(I_0)$ such that $P \subset P_s$, $|P| \leqslant \varepsilon_0^{1+\tau}$ for all $t \in I_0$, and which are maximal with this property. Such a (P, Q, n) is therefore the simple composition of (P_s, Q_s, n_s) with an element $(P', Q', n') \in \mathcal{R}(I_0)$ satisfying, according to (3.12) in Section 3.3

$$C^{-1}\varepsilon_0^{\tau} < |P'| < C\varepsilon_0^{\tau}$$

for all $t \in I_0$. As recalled in the beginning of Section 8.3, for each $t \in I_0$, the number of P' with $|P'| > C^{-1}\varepsilon_0^{\tau}$ is of order $\varepsilon_0^{-\tau d_s^t}$ which is also the order of $\varepsilon_0^{-\tau d_s^0}$. This proves $(SR1)_s$. The estimate $(SR1)_u$ is obtained in the same way.

We turn to $(SR2)'_s$. Let $((P_i, Q_i, n_i))_i$ be a finite family of elements of $\mathcal{R}(I_0)$ with the P_i disjoint and contained in P_s. We have to prove that

$$\sum_i \max_{I_0} |Q_i|^{\rho_u} \leqslant C|Q_s|^{\rho_s}. \tag{$(SR2)'_s$}$$

We can write each (P_i, Q_i, n_i) as the simple composition of (P_s, Q_s, n_s) with an element $(P'_i, Q'_i, n'_i) \in \mathcal{R}(I_0)$. The P'_i are disjoint. We need to have

$$\sum_i \max_{I_0} |Q'_i|^{\rho_u} \leqslant C. \tag{9.39}$$

This will be a consequence from the existence of equilibrium measures for Hölder potentials on regular Cantor sets defined by expansive $C^{1+\alpha}$ maps: fix a parameter $t \in I_0$; choose an horizontal segment in each rectangle of the Markov partition, and let J be their union; the intersection of J with the local stable foliation $W^s(K, R)$ is a regular Cantor set K_s; there exists on K_s a probability measure (a Gibbs state for the appropriate potential) such that, for each $(P, Q, n) \in \mathcal{R}(I_0)$, the measure of the cylinder of K_s defined by P, divided by $|Q|^{d_u^t}$, is uniformly bounded away from 0 and ∞. As the P'_i are disjoint, this proves that, for each $t \in I_0$, we have

$$\sum_i |Q'_i|^{d_u^t} \leqslant C. \tag{9.40}$$

But Proposition 15 in Section 7.4 shows that, for each $(P, Q, n) \in \mathcal{R}(I_0)$

$$\min_{I_0} |Q'_i| \geqslant \left(\max_{I_0} |Q'_i|\right)^{1+C\varepsilon_0}. \tag{9.41}$$

Therefore (9.39) is a consequence of (9.40), (9.41) as

$$\rho_u > d_u^0 + C\varepsilon_0. \tag{9.42}$$

The estimate $(SR2)_u$ is proved in the same way, under the requirement

$$\rho_s > d_s^0 + C\varepsilon_0. \tag{9.43}$$

Let us now prove $(SR3)_s$ and $(SR3)_u$. The only case to consider is $I_\alpha = I_\omega = I_0$. Then $Bi_+(I_0, I_0, I_0; x)$ (resp. $Bi_-(I_0, I_0, I_0; x)$) is the set of $(P, Q, n) \in \mathcal{R}(I_0)$ such that $P \subset P_s$, $Q \subset Q_u$ and $|P| \geqslant x$ (resp. $|Q| \geqslant x$) for some $t \in I_0$. We write (P, Q, n) as a simple composition

$$(P, Q, n) = (P_s, Q_s, n_s) * (P', Q', n') * (P_u, Q_u, n_u)) \tag{9.44}$$

(cf. (5.18) in Section 5.6.4). The same argument than for $(SR1)_s$ above now gives

$$\#Bi_+(I_0, I_0, I_0; x) \leqslant C\left(\frac{x}{\varepsilon_0 |P_u|}\right)^{-(d_s^0 + C\varepsilon_0)}, \tag{9.45}$$

and similarly

$$\#Bi_-(I_0, I_0, I_0; x) \leqslant C\left(\frac{x}{\varepsilon_0 |Q_s|}\right)^{-(d_u^0 + C\varepsilon_0)}. \tag{9.46}$$

Therefore, we obtain $(SR3)_s$ and $(SR3)_u$ as we have:

$$\rho_0 > d_s^0 + C\varepsilon_0, \tag{9.47}$$

$$\rho'_0 > d_u^0 + C\varepsilon_0. \tag{9.48}$$

□

9.5. *The induction step for* (SR1). — In Proposition 28 of Section 8.3, we introduced $d_s^+ = d_s^0 + C\eta\tau^{-1}$. Let also $d_u^+ = d_u^0 + C\eta\tau^{-1}$. The exponent σ in the estimates (SR1) will be defined as

$$\sigma = \min(1 - d_u^+ - (1+\tau)d_s^* - \tau - c\eta\tau^{-1}, 1 - d_s^+ - (1+\tau)d_u^* - \tau - c\eta\tau^{-1}). \tag{9.49}$$

The aim of this Subsection is to prove the following result.

Proposition **32.** — *Assume that the parent interval* $\widetilde{I}$ *is* β*-regular and satisfies one of the two inequalities* (SR1). *Then all candidates* $I \subset \widetilde{I}$ *satisfy the same inequality except perhaps for a proportion not larger than* $C|\widetilde{I}|^{\tau^2}$.

Before proving the Proposition above, we first state and prove

Proposition **33.** — *For any* $(P, Q, n) \in \mathcal{C}_+(\widetilde{I})$, *we have*

$$\#Cr(P) \leqslant C|\widetilde{I}|^{-\tau d_u^+}.$$

Proof. — Let $(P^*, Q^*, n^*) \in \mathcal{R}(\widetilde{I})$ be an element such that $P^* \subset P$ and

$$|P^*| \leqslant \frac{1}{2}|\widetilde{I}|^{(1+\tau)(1-\eta)^{-1}} \tag{9.50}$$

for all $t \in \widetilde{I}$. By Proposition 28 in Section 8.3, there are at most $|\widetilde{I}|^{-\tau d_u^+}$ candidates $I \subset \widetilde{I}$ such that P^* is I-critical.

Let now $I \in Cr(P)$. We will prove that I is within distance $C|\widetilde{I}|^{1+\tau}$ either from the boundary of $\widetilde{I}$, or from a candidate $I' \subset \widetilde{I}$ such that P^* is I′-critical. This implies the estimate of the Proposition.

By definition, there exists $(P_1, Q_1, n_1) \in \mathcal{R}(I)$ such that $P_1 \subset P$, P_1 is I-critical and

$$|P_1|^{1-\eta} \leqslant 2|I| \quad \text{for some } t_1 \in I. \tag{9.51}$$

We take such a (P_1, Q_1, n_1) with P_1 maximal. Then, (P_1, Q_1, n_1) belongs in fact to $\mathcal{R}(\widetilde{I})$: indeed, otherwise, by the structure theorem, the element $(\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1) \in \mathcal{R}(\widetilde{I})$ with smallest $\widetilde{P}_1 \supset P_1$ would be $\widetilde{I}$-bicritical and we would have $|\widetilde{P}_1| < |\widetilde{I}|^{\beta} < 2|I|$ for all $t \in \widetilde{I}$.

As P_1 is I-critical, it is I-special. From Proposition 18 in Section 7.7, we obtain that

$$|P_1|^{1-\eta} \leqslant C|I| \quad \text{for all } t \in I. \tag{9.52}$$

By Corollary 10 in Section 8.1, there exists $(P_0, Q_0, n_0) \in \mathcal{R}(I)$ such that Q_0, P_1 are I-critically related and $|Q_0| < \max(|I|^{\beta}, 2|P_1|)$ for some $t_0 \in I$. In particular, we have

$$|Q_0|^{1-\eta} \leqslant C|I| \quad \text{for some } t_0 \in I. \tag{9.53}$$

We take such a (P_0, Q_0, n_0) with Q_0 maximal. We claim that (P_0, Q_0, n_0) belongs to $\mathcal{R}(\widetilde{I})$ and that Q_0 is $\widetilde{I}$-critical.

Indeed, let $(\widetilde{P}_0, \widetilde{Q}_0, \widetilde{n}_0)$ the element of $\mathcal{R}(\widetilde{I})$ such that $\widetilde{Q}_0$ contains Q_0 and is the smallest with this property.

Assume by contradiction that $\widetilde{Q}_0$ is $\widetilde{I}$-transverse. Let $(P_\alpha, Q_\alpha, n_\alpha)$ be an $\widetilde{I}$-decomposition of P_s such that, for each α, $\widetilde{Q}_0$ and P_α are either $\widetilde{I}$-separated or $\widetilde{I}$-transverse. If $P_1 \subset P_\alpha$ for some α, we would conclude that $\widetilde{Q}_0$ and P_1 are $\widetilde{I}$-separated if $\widetilde{Q}_0$ and P_α are $\widetilde{I}$-separated, and $\widetilde{Q}_0$ and P_1 are $\widetilde{I}$-transverse if $\widetilde{Q}_0$ and P_α are $\widetilde{I}$-transverse. If there exists α such that $P_\alpha \subset P_1$ and $\widetilde{Q}_0 \pitchfork_{\widetilde{I}} P_\alpha$ holds, then $\widetilde{Q}_0 \pitchfork_{\widetilde{I}} P_1$ holds also by Proposition 10 in Section 6.4. In the remaining case, the $P_\alpha \subset P_1$ form an $\widetilde{I}$-decomposition of P_1 and they are all $\widetilde{I}$-separated from $\widetilde{Q}_0$; this imply that P_1 itself is $\widetilde{I}$-separated from $\widetilde{Q}_0$. The contradiction obtained in all cases show that $\widetilde{Q}_0$ is $\widetilde{I}$-critical.

But then, if $Q_0 \neq \widetilde{Q}_0$, the element $(\widetilde{P}_0, \widetilde{Q}_0, \widetilde{n}_0)$ is $\widetilde{I}$-bicritical by the structure theorem and we would have $|\widetilde{Q}_0| < |\widetilde{I}|^\beta$ for all $t \in \widetilde{I}$, contradicting the maximality of Q_0.

Thus we know that both Q_0 and P_1 are $\widetilde{I}$-defined and $\widetilde{I}$-special.In particular, from Proposition 18 in Section 7.7 and (9.53), we have

$$|Q_0|^{1-\eta} \leqslant C|I| \quad \text{for all } t \in \widetilde{I}. \tag{9.54}$$

If we had $\delta_{LR}(Q_0, P_1) > C_0|I|$ for all $t \in I$ and C_0 large enough, we would have $Q_0 \overline{\pitchfork}_I P_1$ from the estimates on $|Q_0|, |P_1|$ above. On the other hand, if we had $\delta_{LR}(Q_0, P_1) < 0$ for all $t \in I$, Q_0 and P_1 would be I-separated. As they are I-critically related, we must have, for some $t \in I$ and $C_0 > 2$

$$0 \leqslant \delta_{LR}(Q_0, P_1) \leqslant C_0|I|. \tag{9.55}$$

Assume that I is not within distance $4C_0|I|$ from the boundary of $\widetilde{I}$. Let J be the $3C_0|I|$-neighborhood of I contained in $\widetilde{I}$. From Corollary 8 in Section 7.6, the quantity $\delta_{LR}(Q_0, P_1)$ will be larger than $2C_0|I|$ at the upper endpoint of J and less than $-C_0|I|$ at the lower endpoint. As P_1 and P^* are both contained in P and $|P| < |I|$ for all $t \in I$, we have, (taking C_0 larger if necessary) that, for all $t \in I$

$$|\delta_{LR}(Q_0, P^*) - \delta_{LR}(Q_0, P_1)| \leqslant C_0|I|. \tag{9.56}$$

Therefore, the value of $\delta_{LR}(Q_0, P^*)$ at the lower endpoint of J is less than 0 and the value at the upper endpoint is at least $C_0|I|$. Then, in view of the estimates on $|Q_0|, |P^*|$ and Lemma 3 in Section 6.6.3, there is a candidate interval I′ intersecting J such that P^* is I′-critical. As explained above, this allows to obtain the conclusion of the proposition. □

Proof of Proposition 32. Let $(\widetilde{P}, \widetilde{Q}, \widetilde{n}) \in \mathcal{C}_+(\widetilde{I})$; we bound the number of $(P, Q, n) \in \mathcal{C}_+(I)$ (for some candidate $I \subset \widetilde{I}$) with $P \subset \widetilde{P}$. Recall that $(P, Q, n) \in \mathcal{R}(\widetilde{I})$ by Lemma 11 in Section 9.1. Let $\widehat{P}$ be the parent of P (if $P \neq \widetilde{P}$). It satisfies $|\widehat{P}| \geqslant |\widetilde{I}|^{(1+\tau)^2}$ for some $t \in \widetilde{I}$. By Corollary 13 in Section 8.3, the number of possible $\widehat{P}$ (for fixed $\widetilde{P}$) is at most $|\widetilde{I}|^{-d_s^*\tau(1+\tau)}$. From Proposition 25 in Section 8.2, each $\widehat{P}$ has at most $|\widetilde{I}|^{-c\eta}$ children. We conclude that the number of possible P for each $\widetilde{P}$ is at most $|\widetilde{I}|^{-d_s^*\tau(1+\tau)-c\eta}$.

We obtain therefore, in view of Lemma 11, Proposition 33 and the induction hypothesis

$$\begin{aligned}
\sum_{I \subset \widetilde{I}} \#\mathcal{C}_+(I) &\leqslant |\widetilde{I}|^{-c\eta-\tau(1+\tau)d_s^*} \sum_{\mathcal{C}_+(\widetilde{I})} \#\mathrm{C}r(\widetilde{P}) \\
&\leqslant \#\mathcal{C}_+(\widetilde{I})|\widetilde{I}|^{-c\eta-\tau(1+\tau)d_s^*} \max \#\mathrm{C}r(\widetilde{P}) \\
&\leqslant C\#\mathcal{C}_+(\widetilde{I})|\widetilde{I}|^{-c\eta-\tau d_u^+-\tau(1+\tau)d_s^*} \\
&\leqslant C\left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^\sigma \varepsilon_0^{-\tau d_s^0} |\widetilde{I}|^{-c\eta-\tau d_u^+-\tau(1+\tau)d_s^*}.
\end{aligned} \tag{9.57}$$

As the total number of candidates is $|\widetilde{\mathrm{I}}|^{-\tau}$, we will have, except for a proportion at most $\mathrm{C}|\widetilde{\mathrm{I}}|^{\tau^2}$ of candidates I

$$\#\mathcal{C}_+(\mathrm{I}) \leqslant \left(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0}\right)^{\sigma} \varepsilon_0^{-\tau d_s^0} |\widetilde{\mathrm{I}}|^{\tau-\tau^2-c\eta-\tau d_u^+-\tau(1+\tau)d_s^*}$$

$$\leqslant \left(\frac{|\mathrm{I}|}{\varepsilon_0}\right)^{\sigma} \varepsilon_0^{-\tau d_s^0} |\widetilde{\mathrm{I}}|^{\tau-\tau^2-c\eta-\tau d_u^+-\tau(1+\tau)d_s^*-\tau\sigma}. \tag{9.58}$$

In order to obtain the required estimate, we want to have

$$\tau - \tau^2 - c\eta - \tau d_u^+ - \tau(1+\tau)d_s^* - \tau\sigma \geqslant 0, \tag{9.59}$$

which amounts to

$$\sigma \leqslant 1 - d_u^+ - (1+\tau)d_s^* - \tau - c\eta\tau^{-1}. \tag{9.60}$$

In the same way, to obtain $(\mathrm{SR1})_u$ for all but a proportion $\mathrm{C}|\widetilde{\mathrm{I}}|^{\tau^2}$ of candidates I from $(\mathrm{SR1})_u$ for $\widetilde{\mathrm{I}}$, we will ask that

$$\sigma \leqslant 1 - d_s^+ - (1+\tau)d_u^* - \tau - c\eta\tau^{-1}. \tag{9.61}$$

In conclusion, we take σ to be the largest number satisfying (9.60), (9.61) above and the proof of the proposition is complete.

9.6. *The induction step for* (SR2)′.

9.6.1. Recall that the exponents ρ_0' (in $(\mathrm{SR3})_u$), ρ_u (in $(\mathrm{SR2})_s'$) and d_u^* are all close to d_u^0 when $\tau \gg \eta \gg \varepsilon_0$ are small. Similarly, ρ_0 (in $(\mathrm{SR3})_s$), ρ_s (in $(\mathrm{SR2})_u'$) and d_s^* are all close to d_s^0. We now assume moreover that

$$\rho_0' - \varepsilon_0^{\tau} > \rho_u > d_u^* + \varepsilon_0^{\tau}, \qquad \rho_0 - \varepsilon_0^{\tau} > \rho_s > d_s^* + \varepsilon_0^{\tau}. \tag{9.62}$$

The exponents σ_s, σ_u in (SR2)′ are defined as

$$\sigma_s := 1 - 3\tau - d_s^+, \qquad \sigma_u := 1 - 3\tau - d_u^+. \tag{9.63}$$

The aim of this subsection is to prove the following result

Proposition **34.** — *Assume that the parent interval* $\widetilde{\mathrm{I}}$ *is strongly regular. Then all candidates* $\mathrm{I} \subset \widetilde{\mathrm{I}}$ *satisfy* $(\mathrm{SR2})_s'$ *and* $(\mathrm{SR2})_u'$ *except perhaps for a proportion not larger than* $\mathrm{C}|\widetilde{\mathrm{I}}|^{\tau^2}$.

Proof. — We first explain the general idea (for $(\mathrm{SR2})_s'$).

Let $(P, Q, n) \in \mathcal{C}_+(\widetilde{I})$. Let $((P_i, Q_i, n_i))_i$ be a finite family of elements of $\mathcal{R}(\widetilde{I})$ with the P_i disjoint and contained in P. As $(SR2)'_s$ is satisfied by $\widetilde{I}$, and P is contained in some P^* with $(P^*, Q^*, n^*) \in \widehat{\mathcal{C}}_+(\widetilde{I})$ by Lemma 11, the supremum over such families of the quantity

$$\sum_i \max_{\widetilde{I}} |Q_i|^{\rho_u}$$

is finite. Denote this supremum by $c(P)$. We have, still from $(SR2)'_s$, that

$$\sum_{\mathcal{C}_+(\widetilde{I})} c(P) \leqslant C|Q_s|^{\rho_u} \left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{\sigma_u}. \tag{9.64}$$

Let now $I \subset \widetilde{I}$ be a candidate interval such that $I \in Cr(P)$. Let $((P_i, Q_i, n_i))_i$ be a finite family of elements of $\mathcal{R}(I)$ with the P_i disjoint and contained in P. Denote by $c(P, I)$ the supremum (in $\mathbf{R} \cup +\infty$) over all such families of the quantity

$$\sum_i \max_{I} |Q_i|^{\rho_u}.$$

We also consider two related quantities $c(P, I, \mathit{old})$ and $c(P, I, \mathit{new})$ obtained by taking the supremum of the same quantity $\sum_i \max_I |Q_i|^{\rho_u}$ over a more restricted set of families: we still ask that the P_i are disjoint and contained in P, but for $c(P, I, \mathit{old})$ we ask moreover that all (P_i, Q_i, n_i) belong to $\mathcal{R}(\widetilde{I})$, and for $c(P, I, \mathit{new})$ that none of them belong to $\mathcal{R}(\widetilde{I})$.

Clearly, we have

$$c(P, I) \leqslant c(P, I, \mathit{old}) + c(P, I, \mathit{new}),$$

and also

$$c(P, I, \mathit{old}) \leqslant c(P). \tag{9.65}$$

On the other hand, by Lemma 11, the candidate I will satisfy $(SR2)'_s$ iff

$$\sum_{\widehat{\mathcal{C}}_+(I)} c(P, I) \leqslant C|Q_s|^{\rho_u} \left(\frac{|I|}{\varepsilon_0}\right)^{\sigma_u}. \tag{9.66}$$

In each term of the sum, we separate the old and new part and will deal successively with these two terms.

9.6.2. The sum over the old parts is easily controlled for most candidates. We have

$$\sum_{I \subset \widetilde{I}} \sum_{\widehat{\mathcal{C}}_+(I)} c(P, I, \mathit{old}) \leqslant \sum_{I \subset \widetilde{I}} \sum_{\widehat{\mathcal{C}}_+(I)} c(P) \quad \text{by (9.72)}$$

$$\leqslant \sum_{\mathcal{C}_+(\widetilde{\mathrm{I}})} c(\mathrm{P})\#\mathrm{C}r(\mathrm{P})$$

$$\leqslant \mathrm{C}|\widetilde{\mathrm{I}}|^{-\tau d_u^+} \sum_{\mathcal{C}_+(\widetilde{\mathrm{I}})} c(\mathrm{P}) \quad \text{by Proposition 33}$$

$$\leqslant \mathrm{C}|\widetilde{\mathrm{I}}|^{-\tau d_u^+} |\mathrm{Q}_s|^{\rho_u} \left(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0}\right)^{\sigma_u}, \quad \text{by (9.71).} \tag{9.67}$$

Consequently, all candidates I except for a proportion at most $|\widetilde{\mathrm{I}}|^{2\tau^2}$ will satisfy

$$\sum_{\widehat{\mathcal{C}}_+(\mathrm{I})} c(\mathrm{P}, \mathrm{I}, old) \leqslant \mathrm{C}|\widetilde{\mathrm{I}}|^{\tau - 2\tau^2 - \tau d_u^+} |\mathrm{Q}_s|^{\rho_u} \left(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0}\right)^{\sigma_u}. \tag{9.68}$$

From the definition $\sigma_u := 1 - 3\tau - d_u^+$ we obtain, for all candidates I except for a proportion at most $|\widetilde{\mathrm{I}}|^{2\tau^2}$

$$\sum_{\widehat{\mathcal{C}}_+(\mathrm{I})} c(\mathrm{P}, \mathrm{I}, old) \leqslant |\mathrm{Q}_s|^{\rho_u} \left(\frac{|\mathrm{I}|}{\varepsilon_0}\right)^{\sigma_u}. \tag{9.69}$$

9.6.3. The sum over the new parts is more complicated.

Let $\mathrm{I} \subset \widetilde{\mathrm{I}}$ be a candidate and let $(\mathrm{P}', \mathrm{Q}', n')$ be an element in $\mathcal{R}(\mathrm{I})$ but not in $\mathcal{R}(\widetilde{\mathrm{I}})$ such that P' is contained in some P with $(\mathrm{P}, \mathrm{Q}, n) \in \widehat{\mathcal{C}}_+(\mathrm{I})$. For a finite family of such elements with disjoint P', we must bound the sum $\mathrm{S} := \sum \max_{\mathrm{I}} |\mathrm{Q}'|^{\rho_u}$.

We apply the structure theorem of Section 6.5. We obtain an integer $k > 0$ and elements $(\mathrm{P}'_l, \mathrm{Q}'_l, n'_l) \in \mathcal{R}(\widetilde{\mathrm{I}})$ for $0 \leqslant l \leqslant k$ such that

- $(\mathrm{P}', \mathrm{Q}', n')$ is obtained from the parabolic composition of the $(\mathrm{P}'_l, \mathrm{Q}'_l, n'_l)$;
- for $0 \leqslant l < k$, Q'_l and P'_{l+1} are $\widetilde{\mathrm{I}}$-critical and $\mathrm{Q}'_l \pitchfork_{\widetilde{\mathrm{I}}} \mathrm{P}'_{l+1}$ does not hold;
- for all $t \in \mathrm{I}$, $|\mathrm{Q}'| \leqslant \mathrm{C}^k |\mathrm{I}|^{-k/2} \prod |\mathrm{Q}'_l|$ (cf. Proposition 13 in Section 6.6.3).

Recall that, by Lemma 10 in Section 9.1, Q'_l and P'_{l+1} are thin $\widetilde{\mathrm{I}}$-critical for $0 \leqslant l < k$. Observe that P'_0 is also thin $\widetilde{\mathrm{I}}$-critical because it is contained in the rectangle P such that $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{C}_+(\widetilde{\mathrm{I}})$ and $\mathrm{P}' \subset \mathrm{P}$.

We claim that P'_k is contained in some $\widetilde{\mathrm{P}}$ with $(\widetilde{\mathrm{P}}, \widetilde{\mathrm{Q}}, \widetilde{n}) \in \widehat{\mathcal{C}}_+(\widetilde{\mathrm{I}})$. As P'_k is thin $\widetilde{\mathrm{I}}$-critical, it is sufficient to prove that a simple descendant $\overline{\mathrm{P}}$ of P'_k with $\max_{\widetilde{\mathrm{I}}} |\overline{\mathrm{P}}| < |\widetilde{\mathrm{I}}|^{1+\tau}$ is $\widetilde{\mathrm{I}}$-critical. If $\overline{\mathrm{P}}$ was $\widetilde{\mathrm{I}}$-transverse, there would exist $(\overline{\mathrm{P}}', \overline{\mathrm{Q}}', \overline{n}')$ with $\overline{\mathrm{Q}}' \cap \mathrm{Q}'_{k-1} \neq \emptyset$ and $\overline{\mathrm{Q}}' \pitchfork_{\widetilde{\mathrm{I}}} \overline{\mathrm{P}}$. As $\mathrm{Q}'_{k-1} \pitchfork_{\mathrm{I}} \mathrm{P}'_k$ also holds, we would deduce by concavity that $\mathrm{Q}'_{k-1} \pitchfork_{\widetilde{\mathrm{I}}} \mathrm{P}'_k$ holds, a contradiction which proves the claim.

Consider now a finite family of such $(\mathrm{P}', \mathrm{Q}', n')$ with disjoint P'. In the sum $\mathrm{S} = \sum \max_{\mathrm{I}} |\mathrm{Q}'|^{\rho_u}$, we first do the partial sum over those $(\mathrm{P}', \mathrm{Q}', n')$ which share the same fixed integer k and the same fixed $(\mathrm{P}'_l, \mathrm{Q}'_l, n'_l)$ for $0 \leqslant l < k$. Except for the fact that the parabolic composition of the (P'_l, Q'_l, n'_l) for $0 \leqslant l \leqslant k$ produces 2^k elements, the disjointness of the P' implies the disjointness of the P'_k. Therefore, from the estimate for $|Q'|$ above, we are able to use $(SR2)'_s$ for $\widetilde{I}$ to bound the partial sum by

$$S(Q'_0, \ldots, Q'_{k-1}) := 2^k \left(C|Q_s|^{\rho_u} \left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{\sigma_u} \right) C^{k\rho_u} |I|^{-k\rho_u/2} \prod_{0 \leqslant l < k} \max_{\widetilde{I}} |Q'_l|^{\rho_u}. \tag{9.70}$$

When we now sum over $Q'_0, \ldots, Q'_{k-1}$, we are led to introduce

$$S'_{bi} := \sum \max_{\widetilde{I}} |Q|^{\rho_u}, \tag{9.71}$$

where the sum is taken over all $(P, Q, n) \in \mathcal{R}(\widetilde{I})$ such that P and Q are thin $\widetilde{I}$-critical. This sum will be estimated below. In terms of S'_{bi} we have for the full some S

$$\begin{aligned} S &\leqslant \sum_{k \geqslant 1} \sum_{Q'_0, \ldots, Q'_{k-1}} S(Q'_0, \ldots, Q'_{k-1}) \\ &\leqslant C|Q_s|^{\rho_u} \left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{\sigma_u} \sum_{k \geqslant 1} (2C^{\rho_u} |I|^{-\rho_u/2} S'_{bi})^k \\ &\leqslant C|Q_s|^{\rho_u} \left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{\sigma_u} |I|^{-\rho_u/2} S'_{bi}, \end{aligned} \tag{9.72}$$

provided $|I|^{-\rho_u/2} S'_{bi} < C^{-1}$ with C large enough.

In the next subsection, we will prove the

Proposition **35.** — *Let θ_u be a number (independent of $\tau \gg \eta \gg \varepsilon_0$) such that $d^0_u/2 < \theta_u < 1 - d^0_s$. Under the hypotheses of the proposition, we have (if $\tau \gg \eta \gg \varepsilon_0$ are small enough)*

$$S'_{bi} \leqslant |\widetilde{I}|^{\theta_u}. \tag{9.73}$$

Remark **11.** — It is easy to check that hypothesis (H4) on d^0_u, d^0_s implies that $d^0_u/2 < 1 - d^0_s$.

The estimate of the proposition is sufficient to get $C|I|^{-\rho_u/2} S'_{bi} < |\widetilde{I}|^{\tau\sigma_u}$ and thus

$$S \leqslant |Q_s|^{\rho_u} \left(\frac{|I|}{\varepsilon_0}\right)^{\sigma_u}. \tag{9.74}$$

The proof of $(SR2)_s$ is therefore complete except for the estimate for S'_{bi}.

Proof of Proposition 35. It is easy to relate S'_{bi} to the cardinality of the sets $Bi_-(\widetilde{I}, \widetilde{I}, \widetilde{I}; x)$. Indeed, we have clearly

$$S'_{bi} \leqslant \sum_{l \geqslant 0} \#Bi_-(\widetilde{I}, \widetilde{I}, \widetilde{I}; 2^{-l}) 2^{(1-l)\rho_u}. \tag{9.75}$$

To estimate the cardinality of $Bi_-(\widetilde{I},\widetilde{I},\widetilde{I};2^{-l})$, we will use $(SR3)_u$ except when l is very large; in this case, the following result will provide a slightly better estimate.

Proposition **36.** — *Assume that* $(SR1)_u$ *holds for* $\widetilde{I}$. *Then, for any candidate interval* I, *any* $x>0$, *the number of* $(P,Q,n)\in\mathcal{R}(I)$ *with* Q *thin* I*-critical and* $\max_I|Q|\geqslant x$ *is at most*

$$C\left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{\sigma}\varepsilon_0^{-\tau d_u^0}\left(\frac{x}{|I|}\right)^{-d_u^*}.$$

Proof. — Let $(P,Q,n)\in\mathcal{R}(I)$ with Q thin I-critical. We claim that $Q\subset\widetilde{Q}$ for some $(\widetilde{P},\widetilde{Q},\widetilde{n})\in\mathcal{C}_-(\widetilde{I})$. Indeed, let $(P',Q',n')\in\mathcal{R}(I)$ such that $Q\subset Q'$, Q' is I-critical, and $\min_I|Q'|^{1-\eta}\leqslant 2|I|$. Replacing if necessary Q' by the smallest $\widetilde{I}$-defined rectangle containing Q', we may assume that $(P',Q',n')\in\mathcal{R}(\widetilde{I})$ (see Corollary 5 in Section 6.6.3). Then, as Q' is $\widetilde{I}$-special, we have from Proposition 18 in Section 7.7 that $\max_{\widetilde{I}}|Q'|^{1-\eta}\leqslant C|I|$. Therefore Q', and also Q is contained in $\widetilde{Q}$ for some $(\widetilde{P},\widetilde{Q},\widetilde{n})\in\mathcal{C}_-(\widetilde{I})$.

In the estimate of the Proposition, we can now assume that $x<|I|$, otherwise there is no (P,Q,n) with the required properties. Let $(\widetilde{P},\widetilde{Q},\widetilde{n})\in\mathcal{C}_-(\widetilde{I})$; we have $\max_{\widetilde{I}}|\widetilde{Q}|\leqslant|I|$. From Corollary 13 in Section 8.3, the number of $(P,Q,n)\in\mathcal{R}(I)$ with Q thin I-critical, $Q\subset\widetilde{Q}$, and $\max_I|Q|\geqslant x$ is at most $(\frac{x}{|I|})^{-d_u^*}$.

As the cardinality of $\mathcal{C}_-(\widetilde{I})$ is at most $C(|\widetilde{I}|/\varepsilon_0)^{\sigma}\varepsilon_0^{-\tau d_u^0}$ by $(SR1)_u$, we obtain the required estimate. □

Corollary **14.** — *One has*

$$\#Bi_-(I,I,I;x)\leqslant B'^{\#}:=C\left(\frac{x}{\varepsilon_0|Q_s|}\right)^{-d_u^*}\left(\frac{|I|}{\varepsilon_0}\right)^{\frac{\sigma}{1+\tau}+d_u^*}\varepsilon_0^{-\tau(d_u^0+\frac{\sigma}{1+\tau})}|Q_s|^{-d_u^*}.$$

The estimate of the Corollary does not take into account that, for the elements $(P,Q,n)\in Bi_-(I,I,I;x)$, P is I-critical. Therefore, it is not surprising that it is worse than $(SR3)_u$ for middle-sized x. Indeed, on one side, the constant $\varepsilon_0^{-\tau(d_u^0+\frac{\sigma}{1+\tau})}|Q_s|^{-d_u^*}$ is large (much larger than $\varepsilon_0^{-A\tau}$). On the other side, comparing the exponents of $(\frac{x}{\varepsilon_0|Q_s|})$ and $(\frac{|I|}{\varepsilon_0})$ in $B'^{\#}, B'_1, B'_0$ when $\tau\gg\eta\gg\varepsilon_0$ are small, one gets

- $-d_u^*$ (close to $-d_u^0$) and $\frac{\sigma}{1+\tau}+d_u^*$ (close to $1-d_s^0$) for $B'^{\#}$;
- $-\rho'_1$ (close to $-\frac{d_u^0}{d_s^0+d_u^0}(2d_s^0+d_u^0-1)>-d_u^0$) and $\sigma_0+\sigma_1$ (close to $1-d_u^0\geqslant 1-d_s^0$) for B'_1;
- $-\rho'_0$ (close to $-d_u^0$) and $2\sigma_0+\sigma_1$ (close to $2-d_u^0-d_s^0>1-d_u^0$) for B'_0.

The only range where $B'^{\#}$ is smaller than $B'=\max(B'_0,B'_1)$ is when x is very small because, although ρ'_0 and d_u^* are both close to d_u^0, we have $\rho'_0>d_u^*$. Forgetting about powers of ε_0^τ, we define $\overline{x}'_{\min}$ (for $\widetilde{\mathrm{I}}$) by

$$(\mathbf{9.76})\qquad \left(\frac{\overline{x}'_{\min}}{\varepsilon_0|\mathrm{Q}_s|}\right)^{-d_u^*}\left(\frac{|\mathrm{I}|}{\varepsilon_0}\right)^{\frac{\sigma}{1+\tau}+d_u^*}|\mathrm{Q}_s|^{-d_u^*}=\left(\frac{\overline{x}'_{\min}}{\varepsilon_0|\mathrm{Q}_s|}\right)^{-\rho'_0}\left(\frac{|\mathrm{I}|}{\varepsilon_0}\right)^{2\sigma_0+\sigma_1},$$

which amounts to

$$(\mathbf{9.77})\qquad \frac{\overline{x}'_{\min}}{\varepsilon_0|\mathrm{Q}_s|}=\left[\left(\frac{|\mathrm{I}|}{\varepsilon_0}\right)^{2\sigma_0+\sigma_1-\frac{\sigma}{1+\tau}-d_u^*}|\mathrm{Q}_s|^{d_u^*}\right]^{\frac{1}{\rho'_0-d_u^*}}.$$

The quantity $\sigma_u^* := 2\sigma_0+\sigma_1-\frac{\sigma}{1+\tau}-d_u^*$ is close to $1-d_u^0$ and the exponent $\frac{1}{\rho'_0-d_u^*}$ is very large.

We now come back to the estimation of S'_{bi}. We divide the sum over l in the right-hand term of (9.82) in three parts:

– $\overline{x}'\geqslant 2^{-l}\geqslant x'_{cr}$.

Here $\overline{x}'=\varepsilon_0^{1+\mathrm{A}\tau/\rho'_1}|\mathrm{Q}_s|(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0})^{\frac{\sigma_0+\sigma_1}{\rho'_1}}$ is defined (in (9.29)) by the relation $\varepsilon_0^{-\mathrm{A}\tau}\mathrm{B}'_1=1$ and satisfies $\overline{x}'<|\widetilde{\mathrm{I}}|^\beta$. The set $\mathrm{B}i_-(\widetilde{\mathrm{I}},\widetilde{\mathrm{I}},\widetilde{\mathrm{I}};x)$ is empty for $x>\overline{x}'$. In the range under consideration, the cardinality of the set $\mathrm{B}i_-(\widetilde{\mathrm{I}},\widetilde{\mathrm{I}},\widetilde{\mathrm{I}};2^{-l})$ is bounded by $\varepsilon_0^{-\mathrm{A}\tau}\mathrm{B}'_1$. We get, as $\rho'_1<\rho_u$

$$(\mathbf{9.78})\qquad \mathrm{S}'_1:=\sum_{\overline{x}'\geqslant 2^{-l}\geqslant x'_{cr}}\#\mathrm{B}i_-(\widetilde{\mathrm{I}},\widetilde{\mathrm{I}},\widetilde{\mathrm{I}};2^{-l})2^{(1-l)\rho_u}\leqslant \mathrm{C}\overline{x}'^{\rho_u}.$$

As $\overline{x}'<|\widetilde{\mathrm{I}}|^\beta$, we have

$$(\mathbf{9.79})\qquad \mathrm{S}'_1\ll|\widetilde{\mathrm{I}}|^{d_u^0}\leqslant|\widetilde{\mathrm{I}}|^{1-d_s^0}.$$

– $x'_{cr}\geqslant 2^{-l}\geqslant\overline{x}'_{\min}$.

In this range, the cardinality of the set $\mathrm{B}i_-(\widetilde{\mathrm{I}},\widetilde{\mathrm{I}},\widetilde{\mathrm{I}};2^{-l})$ is bounded by $\varepsilon_0^{-\mathrm{A}\tau}\mathrm{B}'_0$. As we have $\rho'_0>\rho_u+\varepsilon_0^\tau$, we get

$$\mathrm{S}'_2:=\sum_{x'_{cr}\geqslant 2^{-l}\geqslant\overline{x}'_{\min}}\#\mathrm{B}i_-(\widetilde{\mathrm{I}},\widetilde{\mathrm{I}},\widetilde{\mathrm{I}};2^{-l})2^{(1-l)\rho_u}$$

$$\leqslant \mathrm{C}\varepsilon_0^{-\mathrm{A}\tau}(\rho'_0-\rho_u)^{-1}\left(\frac{\overline{x}'_{\min}}{\varepsilon_0|\mathrm{Q}_s|}\right)^{\rho_u-\rho'_0}(\varepsilon_0|\mathrm{Q}_s|)^{\rho_u}\left(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0}\right)^{2\sigma_0+\sigma_1}$$

$$(\mathbf{9.80})\qquad \leqslant \mathrm{C}\varepsilon_0^{-\tau(\mathrm{A}+1)}(\varepsilon_0|\mathrm{Q}_s|)^{\rho_u}|\mathrm{Q}_s|^{-d_u^*\frac{\rho'_0-\rho_u}{\rho'_0-d_u^*}}\left(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0}\right)^{2\sigma_0+\sigma_1-\sigma_u^*\frac{\rho'_0-\rho_u}{\rho'_0-d_u^*}}.$$

Here, $2\sigma_0+\sigma_1$ and σ_u^* are respectively close to $2-d_u^0-d_s^0$ and $1-d_u^0$, $\frac{\rho'_0-\rho_u}{\rho'_0-d_u^*}$ belongs to $(0,1)$, and ρ_u is close to d_u^0. This allows to obtain, when $\tau\gg\eta\gg\varepsilon_0$ are small enough

$$(\mathbf{9.81})\qquad \mathrm{S}'_2\ll|\widetilde{\mathrm{I}}|^{\theta_u}.$$

– $\overline{x}'_{\min} \geqslant 2^{-l}$.

In this range, we use $B'^{\#}$ to bound the cardinality of the set $Bi_-(\widetilde{I}, \widetilde{I}, \widetilde{I}; 2^{-l})$. As we have $\rho_u > d_u^* + \varepsilon_0^\tau$, we get, using (9.83)

$$S'_3 := \sum_{\overline{x}'_{\min} \geqslant 2^{-l}} \#Bi_-(\widetilde{I}, \widetilde{I}, \widetilde{I}; 2^{-l}) 2^{(1-l)\rho_u}$$

$$\leqslant C\varepsilon_0^{-\tau(d_u^0 + \frac{\sigma}{1+\tau})} (\rho_u - d_u^*)^{-1} \left(\frac{\overline{x}'_{\min}}{\varepsilon_0 |Q_s|} \right)^{\rho_u - \rho'_0} (\varepsilon_0 |Q_s|)^{\rho_u} \left(\frac{|\widetilde{I}|}{\varepsilon_0} \right)^{2\sigma_0 + \sigma_1}$$

$$(\mathbf{9.82}) \qquad \leqslant C\varepsilon_0^{-\tau(1 + d_u^0 + \frac{\sigma}{1+\tau})} (\varepsilon_0 |Q_s|)^{\rho_u} |Q_s|^{-d_u^* \frac{\rho'_0 - \rho_u}{\rho'_0 - d_u^*}} \left(\frac{|\widetilde{I}|}{\varepsilon_0} \right)^{2\sigma_0 + \sigma_1 - \sigma_u^* \frac{\rho'_0 - \rho_u}{\rho'_0 - d_u^*}}.$$

Up to a meaningless power of ε_0^τ, this is the same bound as for S'_2.

Summing the three estimates for S'_1, S'_2, S'_3 give the required estimate for S'_{bi}. The proof of $(SR2)'_s$ is now complete.

Proof of $(SR2)'_u$. — The proof of $(SR2)'_u$ is essentially symmetric to the proof of $(SR2)'_s$, but we have to be somewhat careful because we used (SR3) which is not symmetric.

We divide the sums of the $\max_I |P_i|^{\rho_s}$ (over a finite family of elements of $\mathcal{R}(\widetilde{I})$ with the Q_i disjoint and contained in a Q with $(P, Q, n) \in \widehat{\mathcal{C}}_-(I)$) into an old an a new part as for $(SR2)_s$. The old part is dealt with exactly as above. For the new part, we introduce

$$(\mathbf{9.83}) \qquad S_{bi} := \sum \max_{\widetilde{I}} |P|^{\rho_s},$$

where the sum is taken over all $(P, Q, n) \in \mathcal{R}(\widetilde{I})$ such that P and Q are thin $\widetilde{I}$-critical. We now claim that

$$(\mathbf{9.84}) \qquad S_{bi} \leqslant |\widetilde{I}|^{\theta_s},$$

where θ_s is any fixed constant in $(d_s^0/2, 1 - d_u^0)$ independent of $\tau \gg \eta \gg \varepsilon_0$. The deduction of $(SR2)'_u$ from this estimate is the same as for $(SR2)'_s$. To prove the claim, we proceed as in the proof of Proposition 35. The estimate corresponding to Proposition 36 and its Corollary is now

$$(\mathbf{9.85}) \qquad \#Bi_+(I, I, I; x) \leqslant B^{\#} := C\left(\frac{x}{\varepsilon_0 |P_u|} \right)^{-d_s^*} \left(\frac{|I|}{\varepsilon_0} \right)^{\frac{\sigma}{1+\tau} + d_s^*} \varepsilon_0^{-\tau(d_s^0 + \frac{\sigma}{1+\tau})} |P_u|^{-d_s^*}.$$

The threshold where $B^{\#}$ gets better than $B = \max(B_0, B_1)$ is $\overline{x}_{\min}$ with

$$(\mathbf{9.86}) \qquad \frac{\overline{x}_{\min}}{\varepsilon_0 |P_u|} = \left[\left(\frac{|I|}{\varepsilon_0} \right)^{\sigma_s^*} |P_u|^{d_s^*} \right]^{\frac{1}{\rho_0 - d_s^*}}$$

with $\sigma_s^* := 2\sigma_0 + \sigma_1 - \frac{\sigma}{1+\tau} - d_u^*$ close to $1 - d_s^0$. We have

$$\text{(9.87)}\qquad S_{bi} \leqslant \sum_{l \geqslant 0} \# Bi_+(\widetilde{I}, \widetilde{I}, \widetilde{I}; 2^{-l}) 2^{(1-l)\rho_s},$$

and the right-hand sum is divided into three parts:

– $\overline{x} \geqslant 2^{-l} \geqslant x_{cr}$.

Here $\overline{x} = \varepsilon_0^{1+A\tau/\rho_1} |P_u| (\frac{|\widetilde{I}|}{\varepsilon_0})^{\frac{\sigma_0+\sigma_1}{\rho_1}}$ is defined by the relation $\varepsilon_0^{-A\tau} B_1 = 1$ and satisfies $\overline{x} < |\widetilde{I}|^\beta$. The set $Bi_+(\widetilde{I}, \widetilde{I}, \widetilde{I}; x)$ is empty for $x > \overline{x}$. In the range under consideration, the cardinality of the set $Bi_+(\widetilde{I}, \widetilde{I}, \widetilde{I}; 2^{-l})$ is bounded by $\varepsilon_0^{-A\tau} B_1$. We get, as $\rho_1 < \rho_s$

$$\text{(9.88)}\qquad S_1 := \sum_{C\overline{x} \geqslant 2^{-l} \geqslant x_{cr}} \# Bi_+(\widetilde{I}, \widetilde{I}, \widetilde{I}; 2^{-l}) 2^{(1-l)\rho_s} \leqslant C\overline{x}^{\rho_s} \leqslant |\widetilde{I}|^{1-d_u^0}.$$

– $x_{cr} \geqslant 2^{-l} \geqslant \overline{x}_{\min}$.

In this range, the cardinality of the set $Bi_+(\widetilde{I}, \widetilde{I}, \widetilde{I}; 2^{-l})$ is bounded by $\varepsilon_0^{-A\tau} B_0$. As we have $\rho_0 > \rho_s + \varepsilon_0^\tau$, we get

$$\begin{aligned} S_2 &:= \sum_{x_{cr} \geqslant 2^{-l} \geqslant \overline{x}_{\min}} \# Bi_+(\widetilde{I}, \widetilde{I}, \widetilde{I}; 2^{-l}) 2^{(1-l)\rho_s} \\ &\leqslant C\varepsilon_0^{-A\tau} (\rho_0 - \rho_s)^{-1} \left(\frac{\overline{x}_{\min}}{\varepsilon_0 |P_u|} \right)^{\rho_s - \rho_0} (\varepsilon_0 |P_u|)^{\rho_s} \left(\frac{|\widetilde{I}|}{\varepsilon_0} \right)^{2\sigma_0 + \sigma_1} \\ &\leqslant C\varepsilon_0^{-\tau(A+1)} (\varepsilon_0 |P_u|)^{\rho_s} |P_u|^{-d_s^* \frac{\rho_0 - \rho_s}{\rho_0 - d_s^*}} \left(\frac{|\widetilde{I}|}{\varepsilon_0} \right)^{2\sigma_0 + \sigma_1 - \sigma_s^* \frac{\rho_0 - \rho_s}{\rho_0 - d_s^*}}. \qquad \text{(9.89)} \end{aligned}$$

Here, $2\sigma_0 + \sigma_1$ and σ_s^* are respectively close to $2 - d_u^0 - d_s^0$ and $1 - d_s^0$, $\frac{\rho_0 - \rho_u}{\rho_0 - d_s^*}$ belongs to $(0, 1)$, and ρ_s is close to d_s^0. This allows to obtain

$$\text{(9.90)}\qquad S_2 \ll |\widetilde{I}|^{\theta_s}.$$

– the sum S_3 over $\overline{x}_{\min} \geqslant 2^{-l}$ has the same bound as for S_2, up to a meaningless power of ε_0^τ.

Summing the three estimates for S_1, S_2, S_3 give the required estimate for S_{bi}. The proof of $(SR2)'_u$ is complete. □

9.7. *The induction step for* $(SR3)_s$*: general overview and easy cases.*

9.7.1. *Very small values of* x. — Let $I \subset \widetilde{I}$ be a candidate interval. When x is very small, one obtains directly a (trivial) estimate for the cardinality of $Bi_+(I, I_\alpha, I_\omega; x)$ which turns out to be better than $(SR3)_s$. Indeed, from Corollary 13 in Section 8.3, one certainly has

$$\#\mathrm{B}i_+(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x) \leqslant \mathrm{C}x^{-d_s^*}. \tag{9.91}$$

This very rough estimate is nevertheless better than $(\mathrm{SR3})_s$ when x is very small because $\rho_0 > d_s^*$ (see (9.69) above). More precisely, we define

$$x_{\min} := |\widetilde{\mathrm{I}}|^{\mathrm{C}(\rho_0 - d_s^*)^{-1}}, \tag{9.92}$$

with C large enough. For $x \leqslant x_{\min}$, we have $x^{-d_s^*} \ll \mathrm{B}_0$ for all intervals $\mathrm{I}_\alpha, \mathrm{I}_\omega$ containing I. Therefore we have proved the

Proposition **37.** — *The estimate* $(\mathrm{SR3})_s$ *is satisfied for all candidates* I, *all* $\mathrm{I}_\alpha, \mathrm{I}_\omega \supset \mathrm{I}$, *as soon as* $x \leqslant x_{\min}$.

9.7.2. *Old and new elements.* — Let $\mathrm{I} \subset \widetilde{\mathrm{I}}$ be a candidate interval.

Definition **9.** — *We denote by* $\mathrm{B}i_+^{old}(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x)$ *the set of* $(\mathrm{P}, \mathrm{Q}, n) \in \mathrm{B}i_+(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x)$ *which belong to* $\mathcal{R}(\widetilde{\mathrm{I}})$, *by* $\mathrm{B}i_+^{new}(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x)$ *the set of* $(\mathrm{P}, \mathrm{Q}, n) \in \mathrm{B}i_+(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x)$ *which do not belong to* $\mathcal{R}(\widetilde{\mathrm{I}})$.

We will estimate separately the cardinalities of the two sets. The estimate for $\mathrm{B}i_+^{new}(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x)$ is based on the structure theorem of Section 6.5 and is valid for all candidates $\mathrm{I} \subset \widetilde{\mathrm{I}}$. In Section 9.8, we will prove the

Proposition **38.** — *Assume that* $\widetilde{\mathrm{I}}$ *is strongly regular. There exists a constant* $\sigma_2 > 0$, *depending only on* d_u^0, d_s^0, ω_u, *such that one has, for all candidates* $\mathrm{I} \subset \widetilde{\mathrm{I}}$, *all* $\mathrm{I}_\alpha, \mathrm{I}_\omega \supset \mathrm{I}$, *all* $x \geqslant x_{\min}$

$$\#\mathrm{B}i_+^{new}(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x) \leqslant |\widetilde{\mathrm{I}}|^{\sigma_2}\mathrm{B},$$

where $\varepsilon_0^{-\mathrm{A}\tau}\mathrm{B}$ *is the bound for* $\#\mathrm{B}i_+(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x)$ *predicted by* $(\mathrm{SR3})_s$.

On the other hand, the estimate for $\mathrm{B}i_+^{old}(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x)$ is trivial (from the strong regularity of $\widetilde{\mathrm{I}}$) when both I_α and I_ω contain $\widetilde{\mathrm{I}}$. Indeed, one has in this case

$$\mathrm{B}i_+^{old}(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x) \subset \mathrm{B}i_+(\widetilde{\mathrm{I}}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x) \tag{9.93}$$

(we don't have always equality because the maximum value of $|\mathrm{P}|$ is taken over I in one case and over $\widetilde{\mathrm{I}}$ in the other case). Combining this with Proposition 38, we get

Proposition **39.** — *Assume that* $\widetilde{\mathrm{I}}$ *is strongly regular. Then* $(\mathrm{SR3})_s$ *is satisfied for all candidates* $\mathrm{I} \subset \widetilde{\mathrm{I}}$, *all* $\mathrm{I}_\alpha, \mathrm{I}_\omega \supset \widetilde{\mathrm{I}}$, *all* $x \geqslant x_{\min}$.

Proof. — We have only to observe that the series $\sum_{k\geqslant 0}\varepsilon_k^{\sigma_2}$ (where $\varepsilon_k = \varepsilon_0^{(1+\tau)^k}$ is the length of intervals of generation k), related to the iterated application of Proposition 38 is convergent with sum $\leqslant C\varepsilon_0^{\sigma_2}$. □

9.7.3. *The case* $I = I_\alpha \neq I_\omega$. — Let $I \subset \widetilde{I}$ be a candidate interval. We consider here a set $Bi_+^{old}(I, I, I_\omega; x)$ with $I_\omega \supset \widetilde{I}$. We will estimate the size of this set for most candidates $I \subset \widetilde{I}$, assuming that $\widetilde{I}$ is strongly regular.

Let (P, Q, n) be an element of $Bi_+^{old}(I, I, I_\omega; x)$. As P in thin I-critical, there exists $(P^*, Q^*, n^*) \in \mathcal{R}(I)$ such that $P \subset P^*$, P^* is I-critical and $\min_I |P^*|^{1-\eta} \leqslant 2|I|$. By coherence, $(P^*, Q^*, n^*) \in \mathcal{R}(\widetilde{I})$ and P^* is $\widetilde{I}$-special; then, by Proposition 18 in Section 7.7, we have $\max_{\widetilde{I}} |P^*|^{1-\eta} \leqslant C|I|$. It follows that there exists $(\widetilde{P}, \widetilde{Q}, \widetilde{n}) \in \mathcal{C}_+(\widetilde{I})$ such that $P \subset P^* \subset \widetilde{P}$. Moreover, we have $I \in Cr(\widetilde{P})$ by definition of $Cr(\widetilde{P})$. Observe also that P^* is $\widetilde{I}$-critical, hence P is thin $\widetilde{I}$-critical, and that $\max_{\widetilde{I}} |P| \geqslant \max_I |P| \geqslant x$.

Definition **10.** — *For* $(\widetilde{P}, \widetilde{Q}, \widetilde{n}) \in \mathcal{C}_+(\widetilde{I})$, *we denote by* $Bi_+(\widetilde{P})$ *the set of* $(P, Q, n) \in Bi_+(\widetilde{I}, \widetilde{I}, I_\omega; x)$ *such that* $P \subset \widetilde{P}$.

We have just seen that

$$Bi_+^{old}(I, I, I_\omega; x) \subset \bigsqcup_{\substack{(\widetilde{P},\widetilde{Q},\widetilde{n})\in\mathcal{C}_+(\widetilde{I})\\ I\in Cr(\widetilde{P})}} Bi_+(\widetilde{P}). \tag{9.94}$$

As the sets $Bi_+(\widetilde{P})$ are disjoint, we have

$$\sum_{\mathcal{C}_+(\widetilde{I})} \#Bi_+(\widetilde{P}) \leqslant \#Bi_+(\widetilde{I}, \widetilde{I}, I_\omega; x). \tag{9.95}$$

We now sum over candidates $I \subset \widetilde{I}$ the estimate for the cardinality of $Bi_+^{old}(I, I, I_\omega; x)$ deduced from the inclusion above. We obtain, using Proposition 33 in Section 9.5 and (9.102)

$$\begin{aligned}\sum_{I\subset\widetilde{I}} \#Bi_+^{old}(I, I, I_\omega; x) &\leqslant \sum_{\mathcal{C}_+(\widetilde{I})} \#Bi_+(\widetilde{P})\#Cr(\widetilde{P})\\ &\leqslant C|\widetilde{I}|^{-\tau d_u^+} \sum_{\mathcal{C}_+(\widetilde{I})} \#Bi_+(\widetilde{P})\\ &\leqslant C|\widetilde{I}|^{-\tau d_u^+} \#Bi_+(\widetilde{I}, \widetilde{I}, I_\omega; x). \end{aligned}\tag{9.96}$$

From now on, we assume that the exponents σ_0 (close to $1 - d_s^0$) and σ_1 (close to $d_s^0 - d_u^0$) in (SR3) satisfy

$$\sigma_0 + \sigma_1 \leqslant 1 - 3\tau - d_u^+. \tag{9.97}$$

We also assume that

$$(\mathbf{9.98}) \qquad \rho_0 \geqslant d_s^* + \varepsilon_0^{\tau^2/2}.$$

Proposition **40.** — *Assume that* $\widetilde{\mathrm{I}}$ *is strongly regular. Then, except for a proportion of candidates not greater than* $|\widetilde{\mathrm{I}}|^{\tau^2}$*, the estimate* $(\mathrm{SR3})_s$ *is satisfied by* $\mathrm{B}i_+(\mathrm{I},\mathrm{I},\mathrm{I}_\omega;x)$ *for all* $x \geqslant x_{\min}$ *and all* $\mathrm{I}_\omega \supset \widetilde{\mathrm{I}}$.

Proof. — In view of Proposition 38, we have only to consider $\mathrm{B}i_+^{old}(\mathrm{I},\mathrm{I},\mathrm{I}_\omega;x)$. For fixed $x = 2^{-l} \geqslant x_{\min}$ and fixed $\mathrm{I}_\omega \supset \widetilde{\mathrm{I}}$, it follows from (9.96) that we have

$$(\mathbf{9.99}) \qquad \#\mathrm{B}i_+^{old}(\mathrm{I},\mathrm{I},\mathrm{I}_\omega;x) \leqslant \mathrm{C}|\widetilde{\mathrm{I}}|^{\tau-2\tau^2}|\widetilde{\mathrm{I}}|^{-\tau d_u^+}\#\mathrm{B}i_+(\widetilde{\mathrm{I}},\widetilde{\mathrm{I}},\mathrm{I}_\omega;x),$$

except for a proportion of candidates no greater than $|\widetilde{\mathrm{I}}|^{2\tau^2}$.

Let $x_{cr} = \varepsilon_0|\mathrm{P}_u|(\frac{|\mathrm{I}_\omega|}{\varepsilon_0})^{\frac{\sigma_0}{\rho_0-\rho_1}}$. For $x \leqslant x_{cr}$, we have

$$(\mathbf{9.100}) \qquad \#\mathrm{B}i_+(\widetilde{\mathrm{I}},\widetilde{\mathrm{I}},\mathrm{I}_\omega;x) \leqslant \varepsilon_0^{-\mathrm{A}\tau}\Big(\frac{x}{\varepsilon_0|\mathrm{P}_u|}\Big)^{-\rho_0}\Big(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0}\Big)^{\sigma_0+\sigma_1}\Big(\frac{|\mathrm{I}_\omega|}{\varepsilon_0}\Big)^{\sigma_0},$$

while for $x \geqslant x_{cr}$, we have

$$(\mathbf{9.101}) \qquad \#\mathrm{B}i_+(\widetilde{\mathrm{I}},\widetilde{\mathrm{I}},\mathrm{I}_\omega;x) \leqslant \varepsilon_0^{-\mathrm{A}\tau}\Big(\frac{x}{\varepsilon_0|\mathrm{P}_u|}\Big)^{-\rho_1}\Big(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0}\Big)^{\sigma_0+\sigma_1}.$$

As $\sigma_0+\sigma_1 \leqslant 1-3\tau-d_u^+$, we have

$$(\mathbf{9.102}) \qquad \mathrm{C}|\widetilde{\mathrm{I}}|^{\tau-2\tau^2-\tau d_u^+} \ll |\widetilde{\mathrm{I}}|^{\tau(\sigma_0+\sigma_1)},$$

and therefore, when a candidate I satisfies (9.99), we have

$$(\mathbf{9.103}) \qquad \#\mathrm{B}i_+^{old}(\mathrm{I},\mathrm{I},\mathrm{I}_\omega;x) \ll \varepsilon_0^{-\mathrm{A}\tau}\mathrm{B} = \varepsilon_0^{-\mathrm{A}\tau}\max(\mathrm{B}_0,\mathrm{B}_1).$$

Thus, to get the required estimate, it is sufficient to remove a proportion $\leqslant |\widetilde{\mathrm{I}}|^{2\tau^2}$ of candidates for each $x = 2^{-l} \geqslant x_{\min}$ and each $\mathrm{I}_\omega \supset \widetilde{\mathrm{I}}$. The number of values of l is $\leqslant \mathrm{C}(\rho_0-d_s^*)^{-1} \leqslant \mathrm{C}\varepsilon_0^{-\tau^2/2}\log_2|\widetilde{\mathrm{I}}|$. The number of possibilities for I_ω is the level k of the interval I. From $|\mathrm{I}| = \varepsilon_0^{(1+\tau)^k}$, we have

$$(\mathbf{9.104}) \qquad k = \frac{\log(\frac{\log|\mathrm{I}|}{\log\varepsilon_0})}{\log(1+\tau)}.$$

Now, we have, for ε_0 small enough

$$(\mathbf{9.105}) \qquad \mathrm{C}|\widetilde{\mathrm{I}}|^{2\tau^2}\varepsilon_0^{-\tau^2/2}\log_2|\widetilde{\mathrm{I}}|\frac{\log(\frac{\log|\mathrm{I}|}{\log\varepsilon_0})}{\log(1+\tau)} \ll |\widetilde{\mathrm{I}}|^{\tau^2},$$

and the proof of the proposition is complete. □

9.7.4. *The case* $I = I_\omega \neq I_\alpha$. — The case is essentially symmetric to the case that we have just considered above. We will now assume that (9.98) holds and that

$$\sigma_0 \leqslant 1 - 3\tau - d_s^+. \tag{9.106}$$

Proposition **41.** — *Assume that* $\widetilde{I}$ *is strongly regular. Then, except for a proportion of candidates not greater than* $|\widetilde{I}|^{\tau^2}$, *the estimate* $(SR3)_s$ *is satisfied by* $Bi_+(I, I_\alpha, I; x)$ *for all* $x \geqslant x_{\min}$ *and all* $I_\alpha \supset \widetilde{I}$.

Proof. — The only difference with the proof of Proposition 40 is that we now have, for each $x = 2^{-l} \geqslant x_{\min}$, each $I_\alpha \supset \widetilde{I}$

$$\# Bi_+^{old}(I, I_\alpha, I; x) \leqslant C|\widetilde{I}|^{\tau - 2\tau^2} |\widetilde{I}|^{-\tau d_s^+} \# Bi_+(\widetilde{I}, I_\alpha, \widetilde{I}; x), \tag{9.107}$$

except from a proportion of candidates no greater than $|\widetilde{I}|^{2\tau^2}$, where $Bi_+(\widetilde{I}, I_\alpha, \widetilde{I}; x)$ is now controlled by

$$\# Bi_+(\widetilde{I}, I_\alpha, \widetilde{I}; x) \leqslant \varepsilon_0^{-A\tau} \Big(\frac{x}{\varepsilon_0 |P_u|}\Big)^{-\rho_0} \Big(\frac{|\widetilde{I}|}{\varepsilon_0}\Big)^{\sigma_0} \Big(\frac{|I_\alpha|}{\varepsilon_0}\Big)^{\sigma_0 + \sigma_1}, \tag{9.108}$$

for $x \leqslant x_{cr} = \varepsilon_0 |P_u| (\frac{|I_\alpha|}{\varepsilon_0})^{\frac{\sigma_0}{\rho_0 - \rho_1}}$, and by

$$\# Bi_+(\widetilde{I}, I_\alpha, \widetilde{I}; x) \leqslant \varepsilon_0^{-A\tau} \Big(\frac{x}{\varepsilon_0 |P_u|}\Big)^{-\rho_1} \Big(\frac{|\widetilde{I}|}{\varepsilon_0}\Big)^{\sigma_0} \Big(\frac{|I_\alpha|}{\varepsilon_0}\Big)^{\sigma_1}, \tag{9.109}$$

for $x \geqslant x_{cr}$.

Using $\sigma_0 \leqslant 1 - 3\tau - d_s^+$ allows to conclude as in the proof of Proposition 40. □

9.7.5. *The case* $I = I_\alpha = I_\omega$, x *large.* — Except for the proof of Proposition 38, the only case left in the induction step for $(SR3)_s$ is $I = I_\alpha = I_\omega$, $x \geqslant x_{\min}$. In this case, we have

$$B_0 = \Big(\frac{x}{\varepsilon_0 |P_u|}\Big)^{-\rho_0} \Big(\frac{|I|}{\varepsilon_0}\Big)^{2\sigma_0 + \sigma_1}, \tag{9.110}$$

$$B_1 = \Big(\frac{x}{\varepsilon_0 |P_u|}\Big)^{-\rho_1} \Big(\frac{|I|}{\varepsilon_0}\Big)^{\sigma_0 + \sigma_1}. \tag{9.111}$$

Observe that we have

$$Bi_+^{old}(I, I, I; x) \subset Bi_+^{old}(I, I, \widetilde{I}; x). \tag{9.112}$$

In the proof of Proposition 40, we have shown that for $x = 2^{-l} \geqslant \widetilde{x}_{cr} := \varepsilon_0 |P_u| (\frac{|\widetilde{I}|}{\varepsilon_0})^{\frac{\sigma_0}{\rho_0 - \rho_1}}$, we have

$$\# Bi_+^{old}(I, I, \widetilde{I}; x) \leqslant \varepsilon_0^{-A\tau} B_1, \tag{9.113}$$

except for a proportion $\leqslant |\widetilde{I}|^{2\tau^2}$ of candidates I. We get therefore the required estimate for $Bi_+^{old}(I, I, I; x)$ when $x \geqslant \widetilde{x}_{cr}$.

Proposition **42.** — *Assume that* $\widetilde{I}$ *is strongly regular. Then, except for a proportion of candidates not greater than* $|\widetilde{I}|^{\tau^2}$, *the estimate* $(SR3)_s$ *is satisfied by* $Bi_+(I, I, I; x)$ *for all* $x \geqslant \widetilde{x}_{cr}$, *where*

$$\widetilde{x}_{cr} = \varepsilon_0 |P_u| \left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{\frac{\sigma_0}{\rho_0 - \rho_1}}. \tag{9.114}$$

9.7.6. *The case* $I = I_\alpha = I_\omega$, $x_{\min} \leqslant x \leqslant \widetilde{x}_{cr}$: *idea of the proof.* — This remaining case is by far the hardest! In this range of x, the required bound for the cardinality of $Bi_+(I, I, I; x)$ is given by B_0 (except at the very top of the range); the exponent $2\sigma_0 + \sigma_1$ (close to $2 - d_s^0 - d_u^0$) in (9.110) means that we have to take into account the criticality of both P and Q in the selection process.

Let $I \subset \widetilde{I}$ be a candidate interval and let $(P, Q, n) \in Bi_+^{old}(I, I, I; x)$. The argument at the beginning of Section 9.7.3 shows that there exists $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(\widetilde{I})$ such that $P \subset P_\alpha$ and $I \in Cr(P_\alpha)$. Similarly, there exists $(P_\omega, Q_\omega, n_\omega) \in \mathcal{C}_-(\widetilde{I})$ such that $Q \subset Q_\omega$ and $I \in Cr(Q_\omega)$. We also observe that $(P, Q, n) \in Bi_+(\widetilde{I}, \widetilde{I}, \widetilde{I}; x)$.

Definition **11.** — *For* $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(\widetilde{I})$, $(P_\omega, Q_\omega, n_\omega) \in \mathcal{C}_-(\widetilde{I})$, *we denote by* $Bi_+(P_\alpha, Q_\omega)$ *the set of* $(P, Q, n) \in Bi_+(\widetilde{I}, \widetilde{I}, \widetilde{I}; x)$ *such that* $P \subset P_\alpha$ *and* $Q \subset Q_\omega$.

We have just seen that, for any candidate I, we have

$$Bi_+^{old}(I, I, I; x) \subset \bigsqcup_{\substack{(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(\widetilde{I}) \\ I \in Cr(P_\alpha)}} \bigsqcup_{\substack{(P_\omega, Q_\omega, n_\omega) \in \mathcal{C}_-(\widetilde{I}) \\ I \in Cr(Q_\omega)}} Bi_+(P_\alpha, Q_\omega). \tag{9.115}$$

As the sets $Bi_+(P_\alpha, Q_\omega)$ are disjoint, we have

$$\sum_{\mathcal{C}_+(\widetilde{I})} \sum_{\mathcal{C}_-(\widetilde{I})} \# Bi_+(P_\alpha, Q_\omega) \leqslant \# Bi_+(\widetilde{I}, \widetilde{I}, \widetilde{I}; x). \tag{9.116}$$

If we were able to prove (for a "random" candidate I) that the events $I \in Cr(P_\alpha)$ and $I \in Cr(Q_\omega)$ are roughly "independent", i.e. if the proportion of candidates in $Cr(P_\alpha) \cap Cr(Q_\omega)$ was roughly the product $|\widetilde{I}|^{\tau(2-d_s^+-d_u^+)}$ of the proportions in $Cr(P_\alpha)$ and $Cr(Q_\omega)$ (cf. Proposition 33 in Section 9.5), we would be able to proceed as in the proof of Proposition 40. But this is unfortunately not the case.

Instead, we will use some degree of independence, in the range of x under consideration, 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 assume that

$$\# Bi_+(P_\alpha, Q_\omega) = b_+(P_\alpha) b_-(Q_\omega), \tag{9.117}$$

for some functions b_+, b_- on $\mathcal{C}_+(\widetilde{I})$, $\mathcal{C}_-(\widetilde{I})$, respectively.

From (9.115), we would get

$$\# Bi_+^{old}(I, I, I; x) \leqslant \phi_+(I) \phi_-(I), \tag{9.118}$$

with

$$\phi_+(\mathrm{I}) = \sum_{\mathrm{I}\in \mathrm{C}r(\mathrm{P}_\alpha)} b_+(\mathrm{P}_\alpha), \tag{9.119}$$

$$\phi_-(\mathrm{I}) = \sum_{\mathrm{I}\in \mathrm{C}r(\mathrm{Q}_\omega)} b_-(\mathrm{Q}_\omega). \tag{9.120}$$

We now average *separately* ϕ_+ and ϕ_-. We obtain

$$\sum_{\mathrm{I}} \phi_+(\mathrm{I}) \leqslant \Big(\max_{\mathcal{C}_+(\widetilde{\mathrm{I}})} \#\mathrm{C}r(\mathrm{P}_\alpha)\Big) \sum_{\mathcal{C}_+(\widetilde{\mathrm{I}})} b_+(\mathrm{P}_\alpha), \tag{9.121}$$

$$\sum_{\mathrm{I}} \phi_-(\mathrm{I}) \leqslant \Big(\max_{\mathcal{C}_-(\widetilde{\mathrm{I}})} \#\mathrm{C}r(\mathrm{Q}_\omega)\Big) \sum_{\mathcal{C}_-(\widetilde{\mathrm{I}})} b_-(\mathrm{Q}_\omega), \tag{9.122}$$

where $\mathrm{C}r(\mathrm{P}_\alpha)$ and $\mathrm{C}r(\mathrm{Q}_\omega)$ are estimated by Proposition 33 in Section 9.5. From (9.116), we have

$$\sum_{\mathcal{C}_+(\widetilde{\mathrm{I}})} b_+(\mathrm{P}_\alpha) \sum_{\mathcal{C}_-(\widetilde{\mathrm{I}})} b_-(\mathrm{Q}_\omega) = \sum_{\mathcal{C}_+(\widetilde{\mathrm{I}})\times\mathcal{C}_-(\widetilde{\mathrm{I}})} \#\mathrm{B}i_+(\mathrm{P}_\alpha, \mathrm{Q}_\omega) \leqslant \#\mathrm{B}i_+(\widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}; x). \tag{9.123}$$

It is then sufficient to eliminate candidates for which either ϕ_+ or ϕ_- is much above its average value to obtain $(\mathrm{SR3})_s$ for the remaining intervals.

As (9.117) does not hold, we will find an appropriate substitute as follows.

We will subdivide each class $\mathrm{B}i_+(\mathrm{P}_\alpha, \mathrm{Q}_\omega)$ into subclasses $\mathrm{B}i_+(\mathrm{P}_\alpha, \mathrm{Q}_\omega, \ell)$; the index ℓ runs through a finite large set L dependent on $\widetilde{\mathrm{I}}$ and x but independent on P_α and Q_ω. Moreover, we will have functions $b_+(\mathrm{P}_\alpha, \ell)$, $b_-(\mathrm{Q}_\omega, \ell)$ on $\mathcal{C}_+(\widetilde{\mathrm{I}}) \times \mathrm{L}$, $\mathcal{C}_-(\widetilde{\mathrm{I}}) \times \mathrm{L}$, respectively, such that,

$$\#\mathrm{B}i_+(\mathrm{P}_\alpha, \mathrm{Q}_\omega, \ell) \leqslant b_+(\mathrm{P}_\alpha, \ell), b_-(\mathrm{Q}_\omega, \ell). \tag{9.124}$$

We then set, for each $\ell \in \mathrm{L}$:

$$\phi_{+,\ell}(\mathrm{I}) = \sum_{\mathrm{I}\in \mathrm{C}r(\mathrm{P}_\alpha)} b_+(\mathrm{P}_\alpha, \ell), \tag{9.125}$$

$$\phi_{-,\ell}(\mathrm{I}) = \sum_{\mathrm{I}\in \mathrm{C}r(\mathrm{Q}_\omega)} b_-(\mathrm{Q}_\omega, \ell). \tag{9.126}$$

We average each of these functions to get, in view of Proposition 33 in Section 9.5,

$$\sum_{\mathrm{I}} \phi_{+,\ell}(\mathrm{I}) \leqslant \mathrm{C}|\widetilde{\mathrm{I}}|^{-\tau d_u^+} b_+(\ell), \tag{9.127}$$

$$\sum_{\mathrm{I}} \phi_{-,\ell}(\mathrm{I}) \leqslant \mathrm{C}|\widetilde{\mathrm{I}}|^{-\tau d_s^+} b_-(\ell), \tag{9.128}$$

with

$$b_+(\ell) = \sum_{\mathcal{C}_+(\widetilde{\mathrm{I}})} b_+(\mathrm{P}_\alpha, \ell), \tag{9.129}$$

$$b_-(\ell) = \sum_{\mathcal{C}_-(\widetilde{\mathrm{I}})} b_-(\mathrm{Q}_\omega, \ell). \tag{9.130}$$

For each ℓ, we will have

$$\phi_{+,\ell}(\mathrm{I}) \leqslant |\widetilde{\mathrm{I}}|^{\tau(1-d_u^+-3\tau)} b_+(\ell), \tag{9.131}$$

$$\phi_{-,\ell}(\mathrm{I}) \leqslant |\widetilde{\mathrm{I}}|^{\tau(1-d_s^+-3\tau)} b_-(\ell) \tag{9.132}$$

except for a proportion of candidates not greater than $\mathrm{C}|\widetilde{\mathrm{I}}|^{3\tau^2}$. Set

$$\widehat{\mathrm{B}} = \sum_{\mathrm{L}} b_+(\ell) b_-(\ell). \tag{9.133}$$

Because we need to eliminate candidates for each ℓ, L should not be too large. We will show in Section 9.10 that

$$\#\mathrm{L} \leqslant |\widetilde{\mathrm{I}}|^{-\tau^2}. \tag{9.134}$$

Taking into account that we must eliminate candidates for each $x = 2^{-l} \geqslant x_{\min}$, the total proportion of the failed candidates is at most $|\widetilde{\mathrm{I}}|^{\tau^2}$ (cf. proof of Proposition 40 in Section 9.7.3). On the other hand, for the surviving candidates, the discussion above gives

$$\begin{aligned}
\#\mathrm{B}i_+^{old}(\mathrm{I},\mathrm{I},\mathrm{I};x) &\leqslant \sum_{\mathrm{I}\in \mathrm{C}r(\mathrm{P}_\alpha)\cap \mathrm{C}r(\mathrm{Q}_\omega)} \#\mathrm{B}i_+(\mathrm{P}_\alpha, \mathrm{Q}_\omega)\\
&\leqslant \sum_{\mathrm{L}} \sum_{\mathrm{I}\in \mathrm{C}r(\mathrm{P}_\alpha)} \sum_{\mathrm{I}\in \mathrm{C}r(\mathrm{Q}_\omega)} b_+(\mathrm{P}_\alpha,\ell) b_-(\mathrm{Q}_\omega,\ell)\\
&= \sum_{\mathrm{L}} \phi_{+,\ell}(\mathrm{I})\phi_{-,\ell}(\mathrm{I})\\
&\leqslant |\widetilde{\mathrm{I}}|^{\tau(2-d_s^+-d_u^+-6\tau)}\widehat{\mathrm{B}}.
\end{aligned} \tag{9.135}$$

If we are able to prove that

$$|\widetilde{\mathrm{I}}|^{\tau(2-d_s^+-d_u^+-6\tau)}\widehat{\mathrm{B}} \leqslant \varepsilon_0^{-\mathrm{A}\tau}\mathrm{B}, \tag{9.136}$$

where B is the bound from $(\mathrm{SR3})_s$, we get the required conclusion.

In the next four subsections, we will

- prove Proposition 38 (estimate on the size of $\mathrm{B}i_+^{new}(\mathrm{I},\mathrm{I}_\alpha,\mathrm{I}_\omega;x)$);
- define precisely L and the subclasses $\mathrm{B}i_+(\mathrm{P}_\alpha,\mathrm{Q}_\omega,\ell)$;
- bound the cardinality of L (prove (9.134));
- obtain an appropriate estimate for $\widehat{\mathrm{B}}$.

9.8. *Size of* $\mathrm{B}i_{+}^{new}(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x)$.

9.8.1. The goal of Section 9.8 is to prove Proposition 38 in Section 9.7.2. Let us recall the statement: there exists a constant $\sigma_2 > 0$, depending only on d_u^0, d_s^0, ω_u, such that, if $\widetilde{\mathrm{I}}$ is strongly regular, one has, for all candidates $\mathrm{I} \subset \widetilde{\mathrm{I}}$, all $\mathrm{I}_\alpha, \mathrm{I}_\omega \supset \mathrm{I}$, all $x \geqslant x_{\min}$

$$\#\mathrm{B}i_{+}^{new}(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x) \leqslant |\widetilde{\mathrm{I}}|^{\sigma_2}\mathrm{B}, \tag{9.137}$$

where $\varepsilon_0^{-\mathrm{A}\tau}\mathrm{B}$ is the bound for $\#\mathrm{B}i_{+}(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x)$ predicted by $(\mathrm{SR3})_s$.

It is sufficient to prove this estimate when $\mathrm{I}_\alpha, \mathrm{I}_\omega$ contain $\widetilde{\mathrm{I}}$. Indeed, let $\widetilde{\mathrm{I}}_\alpha = \widetilde{\mathrm{I}} \cup \mathrm{I}_\alpha$, $\widetilde{\mathrm{I}}_\omega = \widetilde{\mathrm{I}} \cup \mathrm{I}_\omega$. We have $\mathrm{B}i_{+}^{new}(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x) \subset \mathrm{B}i_{+}^{new}(\mathrm{I}, \widetilde{\mathrm{I}}_\alpha, \widetilde{\mathrm{I}}_\omega; x)$. On the other hand, the bounds $\varepsilon_0^{-\mathrm{A}\tau}\mathrm{B}$, $\varepsilon_0^{-\mathrm{A}\tau}\widetilde{\mathrm{B}}$ predicted by $(\mathrm{SR3})_s$ for $\#\mathrm{B}i_{+}(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x)$, $\#\mathrm{B}i_{+}(\mathrm{I}, \widetilde{\mathrm{I}}_\alpha, \widetilde{\mathrm{I}}_\omega; x)$ respectively satisfy $\mathrm{B} \geqslant |\widetilde{\mathrm{I}}|^{\mathrm{C}\tau}\widetilde{\mathrm{B}}$. Therefore the estimate (9.144) for $\#\mathrm{B}i_{+}^{new}(\mathrm{I}, \widetilde{\mathrm{I}}_\alpha, \widetilde{\mathrm{I}}_\omega; x)$ imply the estimate for $\#\mathrm{B}i_{+}^{new}(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x)$ with a slightly smaller value of $\sigma_2 > 0$.

We will therefore now assume that both $\mathrm{I}_\alpha, \mathrm{I}_\omega$ contain $\widetilde{\mathrm{I}}$.

9.8.2. Let $\mathrm{I} \subset \widetilde{\mathrm{I}}$ be a candidate interval, let $\mathrm{I}_\alpha, \mathrm{I}_\omega$ be parameter intervals containing $\widetilde{\mathrm{I}}$, let $x = 2^{-l} \geqslant x_{\min}$ and let $(\mathrm{P}, \mathrm{Q}, n) \in \#\mathrm{B}i_{+}(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x)$.

We apply the structure theorem of Section 6.5. We obtain an integer $k > 0$ and elements $(\mathrm{P}_i, \mathrm{Q}_i, n_i) \in \mathcal{R}(\widetilde{\mathrm{I}})$ for $0 \leqslant i \leqslant k$ such that

- $(\mathrm{P}, \mathrm{Q}, n) \in (\mathrm{P}_0, \mathrm{Q}_0, n_0) \boxdot \cdots \boxdot (\mathrm{P}_k, \mathrm{Q}_k, n_k)$;
- for $0 \leqslant i < k$, Q_i and P_{i+1} are $\widetilde{\mathrm{I}}$-critical and $\mathrm{Q}_i \pitchfork_{\widetilde{\mathrm{I}}} \mathrm{P}_{i+1}$ does not hold;
- for all $t \in \mathrm{I}$, $|\mathrm{P}| \leqslant \mathrm{C}^k |\mathrm{I}|^{-k/2} \prod_{0 \leqslant i \leqslant k} |\mathrm{P}_i|$ (cf. Proposition 14 in Section 6.6.3).

For $0 \leqslant i \leqslant k$, denote by $x_i = 2^{-l_i}$ the largest integral negative power of 2 such that $|\mathrm{P}_i| \geqslant x_i$ for some $t \in \widetilde{\mathrm{I}}$.

Lemma **12.** — *We have* $(\mathrm{P}_0, \mathrm{Q}_0, n_0) \in \mathrm{B}i_{+}(\widetilde{\mathrm{I}}, \mathrm{I}_\alpha, \widetilde{\mathrm{I}}; x_0)$, $(\mathrm{P}_k, \mathrm{Q}_k, n_k) \in \mathrm{B}i_{+}(\widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}, \mathrm{I}_\omega; x_k)$, *and* $(\mathrm{P}_i, \mathrm{Q}_i, n_i) \in \mathrm{B}i_{+}(\widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}; x_i)$ *for* $0 < i < k$.

Remark **12.** — This lemma is the reason why we need to consider different levels of criticality for P and Q.

Proof. — For $0 \leqslant i < k$, Q_i and P_{i+1} are thin $\widetilde{\mathrm{I}}$-critical by Lemma 10 in Section 9.1. As P is thin I_α-critical, there exists $(\mathrm{P}', \mathrm{Q}', n') \in \mathcal{R}(\mathrm{I}_\alpha)$ with $\mathrm{P} \subset \mathrm{P}'$ and $\min_{\mathrm{I}_\alpha} |\mathrm{P}'|^{1-\eta} \leqslant 2|\mathrm{I}_\alpha|$. As $\mathrm{I}_\alpha \supset \widetilde{\mathrm{I}}$ and P_0 is the thinnest $\widetilde{\mathrm{I}}$-defined rectangle containing P, one has $\mathrm{P}_0 \subset \mathrm{P}'$ and P_0 is thin I_α-critical. Similarly, Q_k is thin I_ω-critical. In view of the definition of the x_i, the proof of the lemma is complete. □

From the definition of the x_i and the estimate for $|\mathrm{P}|$ above, we get

$$x \leqslant \mathrm{C}^k |\mathrm{I}|^{-\frac{k}{2}} \prod_0^k x_i. \tag{9.138}$$

Let us write

$$\#(x_0) := \#\mathrm{B}i_+(\widetilde{\mathrm{I}}, \mathrm{I}_\alpha, \widetilde{\mathrm{I}}; x_0),$$
$$\#(x_k) := \#\mathrm{B}i_+(\widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}, \mathrm{I}_\omega; x_k),$$
$$\#(x_i) := \#\mathrm{B}i_+(\widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}; x_i) \quad \text{for } 0 < i < k.$$

Then, as each parabolic composition produces two elements, we have

$$\#\mathrm{B}i_+^{new}(\mathrm{I}, \mathrm{I}_\alpha, \mathrm{I}_\omega; x) \leqslant \sum_{k>0} 2^k \sum_{x_0,\dots,x_k} \prod_0^k \#(x_i) \tag{9.139}$$

where the x_i in the sum are of the form 2^{-l_i} and must satisfy (9.138).

The term $\#(x_i)$ is estimated by the induction hypothesis $(\mathrm{SR3})_s$ for $\widetilde{\mathrm{I}}$. The bound for $\#(x_i)$ has a phase transition at a threshold $x_{i,cr}$, with (cf. (9.16))

$$\begin{aligned} x_{0,cr} &= \varepsilon_0 |\mathrm{P}_u| \left(\frac{|\mathrm{I}_\alpha|}{\varepsilon_0}\right)^{\frac{\sigma_0}{\rho_0-\rho_1}}, \\ x_{k,cr} &= \varepsilon_0 |\mathrm{P}_u| \left(\frac{|\mathrm{I}_\omega|}{\varepsilon_0}\right)^{\frac{\sigma_0}{\rho_0-\rho_1}}, \\ x_{i,cr} &= \varepsilon_0 |\mathrm{P}_u| \left(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0}\right)^{\frac{\sigma_0}{\rho_0-\rho_1}}, \quad \text{for } 0 < i < k. \end{aligned} \tag{9.140}$$

We divide the sum in the right-hand part of (9.139) into two parts. In the first sum, denoted by S_l, we put the terms for which every x_i is above the threshold $x_{i,cr}$. In the second sum, denoted by S_s, we put the terms for which at least one of the x_i is below $x_{i,cr}$.

9.8.3. *Terms with all x_i large.* — Let us consider a term in S_l. All $\#(x_i)$ are bounded by $\varepsilon_0^{-\mathrm{A}\tau}\mathrm{B}_1$ and we have

$$\prod_0^k \#(x_i) \leqslant \varepsilon_0^{-(k+1)\mathrm{A}\tau} \left(\prod_0^k x_i\right)^{-\rho_1} (\varepsilon_0|\mathrm{P}_u|)^{(k+1)\rho_1} \left(\frac{|\mathrm{I}_\alpha|}{\varepsilon_0}\right)^{\sigma_1} \left(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0}\right)^{(k+1)\sigma_0+k\sigma_1}. \tag{9.141}$$

In view of (9.138), the right-hand side is bounded by

$$\mathrm{C}\left(\frac{x}{\varepsilon_0|\mathrm{P}_u|}\right)^{-\rho_1} \left(\frac{|\mathrm{I}_\alpha|}{\varepsilon_0}\right)^{\sigma_1} \left(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0}\right)^{\sigma_0} \mathrm{Z}^k, \tag{9.142}$$

with

$$\mathrm{Z} := \mathrm{C}\varepsilon_0^{-2\mathrm{A}\tau}\left(\varepsilon_0|\mathrm{P}_u||\mathrm{I}|^{-\frac{1}{2}}\right)^{\rho_1} \left(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0}\right)^{\sigma_0+\sigma_1}. \tag{9.143}$$

For all $0 \leqslant i \leqslant k$, from (9.140) and $C^{-1}\varepsilon_0^{\omega_u} \leqslant |P_u| \leqslant C^{-1}\varepsilon_0^{\omega_u}$ (cf. (9.17) in Section 5.6.4), we have $x_{i,cr} > |\widetilde{I}|^C$ for some large enough C. Therefore, for each $k > 0$, the number of terms in S_l is at most $(C|\log_2 |\widetilde{I}||)^{k+1} \leqslant (C|\log_2 |\widetilde{I}||)^{2k}$.

We thus obtain for S_l the bound

$$S_l \leqslant C\left(\frac{x}{\varepsilon_0|P_u|}\right)^{-\rho_1}\left(\frac{|I_\alpha|}{\varepsilon_0}\right)^{\sigma_1}\left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{\sigma_0}\sum_{k>0} Z_1^k, \tag{9.144}$$

with

$$Z_1 := 2(C|\log_2 |\widetilde{I}||)^2 Z. \tag{9.145}$$

In view of (9.143), we have

$$Z_1 \ll \left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{\sigma_0+\sigma_1-\rho_1} |\widetilde{I}|^{\frac{1}{3}\rho_1}, \tag{9.146}$$

where the exponent $\sigma_0 + \sigma_1 - \rho_1$ is strictly positive from the hypothesis (H4), as we have seen in Section 9.3 (cf. (9.24)).

As $Z_1 < \frac{1}{2}$, we obtain from (9.144) that

$$\begin{aligned} S_l &\leqslant \frac{1}{2}\left(\frac{x}{\varepsilon_0|P_u|}\right)^{-\rho_1}\left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{2\sigma_0+\sigma_1-\rho_1}\left(\frac{|I_\alpha|}{\varepsilon_0}\right)^{\sigma_1}|\widetilde{I}|^{\frac{1}{3}\rho_1} \\ &\leqslant \frac{1}{2}\left(\frac{x}{\varepsilon_0|P_u|}\right)^{-\rho_1}\left(\min\frac{|I_\alpha|}{\varepsilon_0},\frac{|I_\omega|}{\varepsilon_0}\right)^{\sigma_0}\left(\frac{|I_\alpha|}{\varepsilon_0}\right)^{\sigma_1}|\widetilde{I}|^{\frac{1}{3}\rho_1} \\ &= \frac{1}{2}B_1|\widetilde{I}|^{\frac{1}{3}\rho_1}, \end{aligned} \tag{9.147}$$

where we have used that both I_α, I_ω contain $\widetilde{I}$ and $\sigma_0 + \sigma_1 - \rho_1 > 0$.

9.8.4. *Terms with some x_i small.* — Consider a term in the sum S_s. Let J be the non-empty subset of indices $i \in \{0, \ldots, k\}$ 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 (9.138)

$$\prod_J x_i^{-\rho_0} \prod_{J^c} x_i^{-\rho_1} \leqslant \left(C^{-k}|I|^{\frac{k}{2}}x\right)^{-\rho_0}. \tag{9.148}$$

As we also have $\widetilde{I} \subset I_\alpha$, $\widetilde{I} \subset I_\omega$, we obtain

$$\prod_i \#(x_i) \leqslant C\varepsilon_0^{-A\tau}\left(\frac{x}{\varepsilon_0|P_u|}\right)^{-\rho_0}\left(\frac{|I_\alpha|}{\varepsilon_0}\right)^{\sigma_0+\sigma_1}\left(\frac{|I_\omega|}{\varepsilon_0}\right)^{\sigma_0} Y_0^{j-1} Y_1^k, \tag{9.149}$$

with

$$(\mathbf{9.150}) \qquad Y_0 = (\varepsilon_0|P_u|)^{\rho_0-\rho_1}\Big(\frac{|\widetilde{I}|}{\varepsilon_0}\Big)^{\sigma_0},$$

$$(\mathbf{9.151}) \qquad Y_1 = C\varepsilon_0^{-A\tau}|I|^{-\rho_0/2}(\varepsilon_0|P_u|)^{\rho_1}\Big(\frac{|\widetilde{I}|}{\varepsilon_0}\Big)^{\sigma_0+\sigma_1}.$$

We have

$$(\mathbf{9.152}) \qquad Y_0 < 1,$$

and, as $C^{-1}\varepsilon_0^{\omega_u} \leqslant |P_u| \leqslant C^{-1}\varepsilon_0^{\omega_u}$ (cf. (5.17)), we can rewrite Y_1 as

$$(\mathbf{9.153}) \qquad Y_1 = C\Big(\frac{\widetilde{I}}{\varepsilon_0}\Big)^{\sigma_0+\sigma_1-\frac{1}{2}\rho_0(1+\tau)}\varepsilon_0^{\rho_1-\frac{1}{2}\rho_0(1+\tau)+\rho_1\omega_u-A\tau}.$$

The exponent $\sigma_0+\sigma_1-\frac{1}{2}\rho_0(1+\tau)$ is close to $1-d_u^0-1/2d_s^0$ when $\tau \gg \eta \gg \varepsilon_0$ are small. We have $1-d_u^0-1/2d_s^0 > 0$ from (H4) (see also the remark after Proposition 35). The exponent $\rho_1-\frac{1}{2}\rho_0(1+\tau)$ is close to $\frac{1}{2}d_s^0(3d_s^0+d_u^0-2)(d_s^0+d_u^0)^{-1}$ when $\tau \gg \eta \gg \varepsilon_0$ are small. From $d_s^0+d_u^0 \geqslant 1$, $d_s^0 \geqslant d_u^0$, we have $3d_s^0+d_u^0-2 \geqslant 0$ (with equality iff $d_u^0 = d_s^0 = 1/2$). On the other hand, the exponent $\rho_1\omega_u$ is close to $\omega_u d_s^0(2d_s^0+d_u^0-1)(d_s^0+d_u^0)^{-1} > 0$ when $\tau \gg \eta \gg \varepsilon_0$ are small. All this means that we can find a number $\sigma_2 > 0$, depending only on d_u^0, d_s^0, ω_u (and only on d_u^0, d_s^0 when $(d_u^0, d_s^0) \neq (1/2, 1/2)$) such that

$$(\mathbf{9.154}) \qquad Y_1 \leqslant |\widetilde{I}|^{2\sigma_2}.$$

Consider now, for fixed k, the number of terms in S_s. From Proposition 14 in Section 6.6.3, we have $x_i \leqslant |\widetilde{I}|$ for each i; then (9.138) and $x \geqslant x_{\min}$ imply that $x_i > x_{\min}$ for each i. As $x_{\min} = |\widetilde{I}|^{C(\rho_0-d_s^*)^{-1}}$, the number of terms is bounded by

$$(\mathbf{9.155}) \qquad \Big(C(\rho_0-d_s^*)^{-1}\log_2|\widetilde{I}|^{-1}\Big)^{k+1}.$$

Using $k+1 \leqslant 2k$ for $k>0$ and (9.149), (9.152), we obtain the following bound for S_s

$$(\mathbf{9.156}) \qquad S_s \leqslant \Big(\frac{x}{\varepsilon_0|P_u|}\Big)^{-\rho_0}\Big(\frac{|I_\alpha|}{\varepsilon_0}\Big)^{\sigma_0+\sigma_1}\Big(\frac{|I_\omega|}{\varepsilon_0}\Big)^{\sigma_0}\sum_{k>0} Y^k$$

with

$$(\mathbf{9.157}) \qquad Y = 2C(\rho_0-d_s^*)^{-2}(\log|\widetilde{I}|^{-1})^2 Y_1.$$

From (9.98) in Section 9.7.3, we have $(\rho_0-d_s^*)^{-2} \leqslant \varepsilon_0^{-\tau^2} \leqslant |\widetilde{I}|^{-\tau^2}$. We obtain therefore

$$(\mathbf{9.158}) \qquad \sum_{k>0} Y^k < \frac{1}{2}|\widetilde{I}|^{\sigma_2}.$$

We have thus shown that

$$(\mathbf{9.159}) \qquad S_s \leqslant \frac{1}{2} B_0 |\widetilde{I}|^{\sigma_2}.$$

Taking σ_2 smaller if necessary, we can assume that

$$(\mathbf{9.160}) \qquad 0 < \sigma_2 \leqslant \frac{1}{3}\rho_1.$$

The estimates (9.147) and (9.159) complete the proof of Proposition 38.

9.9. *Subclasses of bicritical elements.*

9.9.1. *Complement to the structure theorem.* — The following result will be used several times in the construction of the classes of bicritical elements.

Proposition **43.** — *Let* $I \subset \widetilde{I}$ *be a candidate interval. Let* (P, Q, n), $(P', Q', n') \in \mathcal{R}(I)$ *with* $P \subset P'$. *Assume that* (P, Q, n) *does not belong to* $\mathcal{R}(\widetilde{I})$. *Let* $k > 0$ *be the integer and* $(P_i, Q_i, n_i) \in \mathcal{R}(\widetilde{I})$ *for* $0 \leqslant i \leqslant k$ *be the elements given by the structure theorem.*

1. *If* (P', Q', n') *belongs to* $\mathcal{R}(\widetilde{I})$, *we have* $P' \supset P_0$.
2. *If* (P', Q', n') *does not belong to* $\mathcal{R}(\widetilde{I})$, *there exists* $0 < j \leqslant k$ *and* $(P'_j, Q'_j, n'_j) \in \mathcal{R}(\widetilde{I})$ *with* $P'_j \supset P_j$ *such that*

$$(P', Q', n') \in (P_0, Q_0, n_0) \boxempty \cdots \boxempty (P_{j-1}, Q_{j-1}, n_{j-1}) \boxempty (P'_j, Q'_j, n'_j)$$

is the decomposition associated to (P', Q', n') *by the structure theorem.*

Remark **13.** — Taking $j = 0$ and $(P'_0, Q'_0, n'_0) = (P', Q', n')$, the first case can be considered as a special case of the second case.

Proof. — The first case follows from the fact that P_0 is the thinnest $\widetilde{I}$-defined rectangle containing P.

We now assume that (P', Q', n') does not belong to $\mathcal{R}(\widetilde{I})$. By coherence, P_0 is also the thinnest $\widetilde{I}$-defined rectangle containing P'. Therefore, (P_0, Q_0, n_0) is also the first element in the decomposition associated to (P', Q', n') by the structure theorem.

Let $(\widehat{P}, \widehat{Q}, \widehat{n}) \in (P_1, Q_1, n_1) \boxempty \cdots \boxempty (P_k, Q_k, n_k)$ be the element such that

$$(P, Q, n) \in (P_0, Q_0, n_0) \boxempty (\widehat{P}, \widehat{Q}, \widehat{n}).$$

We claim that, when $k > 1$, i.e. when $(\widehat{P}, \widehat{Q}, \widehat{n})$ does not belong to $\mathcal{R}(\widetilde{I})$,

$$(\widehat{P}, \widehat{Q}, \widehat{n}) \in (P_1, Q_1, n_1) \boxempty \cdots \boxempty (P_k, Q_k, n_k)$$

is the decomposition associated to $(\widehat{P}, \widehat{Q}, \widehat{n})$ by the structure theorem. Indeed, this follows easily from the characterization of the (P_i, Q_i, n_i) in terms of maximal $\widetilde{I}$-intervals (see Lemma 2 in Section 6.5.2).

Define similarly $(\widehat{P}', \widehat{Q}', \widehat{n}')$ such that

$$(P', Q', n') \in (P_0, Q_0, n_0) \,\square\, (\widehat{P}', \widehat{Q}', \widehat{n}').$$

As $P \subset P'$, we have $\widehat{P} \subset \widehat{P}'$. If $k = 1$, we have $(\widehat{P}, \widehat{Q}, \widehat{n}) = (P_1, Q_1, n_1)$, we take $j = 1$, $(P'_1, Q'_1, n'_1) = (\widehat{P}', \widehat{Q}', \widehat{n}')$ to satisfy the conclusions of the proposition. If $k > 1$, the claim above allows also to conclude by induction on k. □

9.9.2. *Bound elements.*

Definition **12.** — *Let* $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(\widetilde{I})$, $(P_\omega, Q_\omega, n_\omega) \in \mathcal{C}_-(\widetilde{I})$. *An element* $(P, Q, n) \in Bi_+(P_\alpha, Q_\omega)$ *is bound if* $n \leqslant n_\alpha + n_\omega$. *Otherwise, we say that* (P, Q, n) *is free. We will denote by* $Bi_+(P_\alpha, Q_\alpha, \Diamond)$ *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* $x_{\min} \leqslant x \leqslant \widetilde{x}_{cr}$. *When* $x \ll \widetilde{x}_{cr}$, *most elements are free. If we would allow* $x \gg \widetilde{x}_{cr}$, *on the opposite, most elements would be bound.*

Proposition **44.** — *For any* $(P_\alpha, Q_\alpha, n_\alpha), (P_\omega, Q_\omega, n_\omega) \in \mathcal{R}(\widetilde{I})$, *and any* $n \leqslant n_\alpha + n_\omega$, *there is at most one element* $(P, Q, n) \in \mathcal{R}(\widetilde{I})$ *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 $\widetilde{I}$ is the starting interval I_0, no parabolic composition is involved and the result follows from usual symbolic dynamics: as $n \leqslant n_\alpha + n_\omega$, 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 $\widetilde{I}$. Denote by $\widetilde{I}_1$ the parent interval of $\widetilde{I}$.

Assume first that both $(P_\alpha, Q_\alpha, n_\alpha)$ and $(P_\omega, Q_\omega, n_\omega)$ belong to $\mathcal{R}(\widetilde{I}_1)$. We claim that any bound element also belongs to $\mathcal{R}(\widetilde{I}_1)$, 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 $\mathcal{R}(\widetilde{I}_1)$, we apply the structure theorem: it gives elements $(P_0, Q_0, n_0), (P_k, Q_k, n_k) \in \mathcal{R}(\widetilde{I}_1)$ such that P_0 is the thinnest rectangle containing P defined over $\widetilde{I}_1$, Q_k is the thinnest rectangle containing Q defined over $\widetilde{I}_1$, and $n \geqslant n_0 + n_k + N_0$. Therefore, $n_0 \geqslant n_\alpha$, $n_k \geqslant n_\omega$ and $n > n_\alpha + n_\omega$.

We now consider the case when, for instance, $(P_\alpha, Q_\alpha, n_\alpha)$ does not belong to $\mathcal{R}(\widetilde{I}_1)$. We now apply the structure theorem to $(P_\alpha, Q_\alpha, n_\alpha)$ and also to an element $(P, Q, n) \in \mathcal{R}(\widetilde{I})$ with $P \subset P_\alpha$, $Q \subset Q_\omega$, $n \leqslant n_\alpha + n_\omega$. From Proposition 43, we obtain integers $0 < j \leqslant k$, elements $(P_i, Q_i, n_i) \in \mathcal{R}(\widetilde{I}_1)$ for $0 \leqslant i \leqslant k$ such that

$$(P, Q, n) \in (P_0, Q_0, n_0) \,\square \cdots \square\, (P_k, Q_k, n_k) \tag{9.161}$$

and also $(\widetilde{P}_j, \widetilde{Q}_j, \widetilde{n}_j) \in \mathcal{R}(\widetilde{I}_1)$ such that $\widetilde{P}_j \supset P_j$ and

$$(P_\alpha, Q_\alpha, n_\alpha) \in (P_0, Q_0, n_0) \,\square \cdots \square\, (P_{j-1}, Q_{j-1}, n_{j-1}) \,\square\, (\widetilde{P}_j, \widetilde{Q}_j, \widetilde{n}_j). \tag{9.162}$$

Similarly, there exists m with $0 \leqslant m \leqslant k$ and $(\widetilde{\mathrm{P}}'_m, \widetilde{\mathrm{Q}}'_m, \widetilde{n}'_m)$ in $\mathcal{R}(\widetilde{\mathrm{I}}_1)$ such that $\widetilde{\mathrm{Q}}'_m \supset \mathrm{Q}_m$ and

- either $m = k$, $(\widetilde{\mathrm{P}}'_m, \widetilde{\mathrm{Q}}'_m, \widetilde{n}'_m) = (\mathrm{P}_\omega, \mathrm{Q}_\omega, n_\omega)$;
- or $m < k$ and we have,

$$(\mathbf{9.163})\qquad (\mathrm{P}_\omega, \mathrm{Q}_\omega, n_\omega) \in (\widetilde{\mathrm{P}}'_m, \widetilde{\mathrm{Q}}'_m, \widetilde{n}'_m) \square \cdots \square (\mathrm{P}_k, \mathrm{Q}_k, n_k).$$

The relations $n \leqslant n_\alpha + n_\omega$, $n_j \geqslant \widetilde{n}_j$, $n_m \geqslant \widetilde{n}'_m$ imply that $m \leqslant j$.

If $m < j$, the sequence $(\mathrm{P}_i, \mathrm{Q}_i, n_i)$ 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 $(\mathrm{P}_i, \mathrm{Q}_i, n_i)$ for $i \neq m = j$ is determined by P_α, Q_ω; but we also have $\mathrm{P}_j \subset \widetilde{\mathrm{P}}_j$, $\mathrm{Q}_j = \mathrm{Q}_m \subset \widetilde{\mathrm{Q}}_m$ and $n_j \leqslant \widetilde{n}_j + \widetilde{n}'_m$, so by the induction hypothesis $(\mathrm{P}_j, \mathrm{Q}_j, n_j)$ is also determined by P_α, Q_ω and n. 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. □

Recall that, by Proposition 12 in Section 6.6.2, we have

$$(\mathbf{9.164})\qquad \max_{\widetilde{\mathrm{I}}} |\mathrm{P}| \leqslant \mathrm{C}_0 \exp(-n^\gamma)$$

for any $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}(\widetilde{\mathrm{I}})$, with $\gamma = \log \frac{3}{2} / \log 2 > \frac{1}{2}$.

If $(\mathrm{P}, \mathrm{Q}, n) \in \mathrm{B}i_+(\mathrm{P}_\alpha, \mathrm{Q}_\omega, \lozenge)$, we have $\max_{\widetilde{\mathrm{I}}} |\mathrm{P}| \geqslant x \geqslant x_{\min} = |\widetilde{\mathrm{I}}|^{\mathrm{C}(\rho_0 - d_s^*)^{-1}}$, and therefore

$$(\mathbf{9.165})\qquad n \leqslant (\rho_0 - d_s^*)^{-2} (\log |\widetilde{\mathrm{I}}|)^2.$$

We thus shall define

$$(\mathbf{9.166})\qquad b_+(\mathrm{P}_\alpha, \lozenge) = b_-(\mathrm{Q}_\omega, \lozenge) = (\rho_0 - d_s^*)^{-1} \log |\widetilde{\mathrm{I}}|^{-1},$$

and we will indeed have, from Proposition 44,

$$(\mathbf{9.167})\qquad \#\mathrm{B}i_+(\mathrm{P}_\alpha, \mathrm{Q}_\omega, \lozenge) \leqslant b_+(\mathrm{P}_\alpha, \lozenge) b_-(\mathrm{Q}_\omega, \lozenge).$$

9.9.3. *Decomposition of a free element.* — Let $(\mathrm{P}_\alpha, \mathrm{Q}_\alpha, n_\alpha) \in \mathcal{C}_+(\widetilde{\mathrm{I}})$, $(\mathrm{P}_\omega, \mathrm{Q}_\omega, n_\omega) \in \mathcal{C}_-(\widetilde{\mathrm{I}})$ and let $(\mathrm{P}, \mathrm{Q}, n) \in \mathrm{B}i_+(\mathrm{P}_\alpha, \mathrm{Q}_\omega)$ be a *free* element. We will analyze with respect to the structure theorem the way in which $(\mathrm{P}, \mathrm{Q}, n)$, $(\mathrm{P}_\alpha, \mathrm{Q}_\alpha, n_\alpha)$, $(\mathrm{P}_\omega, \mathrm{Q}_\omega, n_\omega)$ have been created. This will allow us in the sequel to define various subclasses of free elements.

Denote by $\widehat{\mathrm{I}}_0$ the largest parameter interval such that $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}(\widehat{\mathrm{I}}_0)$. Elements $(\mathrm{P}, \mathrm{Q}, n)$ for which $\widehat{\mathrm{I}}_0$ is the starting interval I_0 are said to have depth 0. They form a first free subclass of $\mathrm{B}i_+(\mathrm{P}_\alpha, \mathrm{Q}_\omega)$ denoted by $\mathrm{B}i_+(\mathrm{P}_\alpha, \mathrm{Q}_\omega, 0)$.

We now assume that $\widehat{I}_0 \neq I_0$ and denote by $\widetilde{I}_0$ the parent interval of $\widehat{I}_0$. We apply the structure theorem. We obtain an integer $k > 0$, elements $(P_0, Q_0, n_0), \dots, (P_k, Q_k, n_k)$ in $\mathcal{R}(\widetilde{I}_0)$ such that

$$(P, Q, n) \in (P_0, Q_0, n_0) \boxdot \cdots \boxdot (P_k, Q_k, n_k). \tag{9.168}$$

By Proposition 43, we find $0 \leqslant j \leqslant k$ and $(\widetilde{P}_j, \widetilde{Q}_j, \widetilde{n}_j) \in \mathcal{R}(\widetilde{I}_0)$ such that either $j = 0$, $(\widetilde{P}_j, \widetilde{Q}_j, \widetilde{n}_j) = (P_\alpha, Q_\alpha, n_\alpha)$ (if $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{R}(\widetilde{I}_0)$) or $j > 0$ and

$$(P_\alpha, Q_\alpha, n_\alpha) \in (P_0, Q_0, n_0) \boxdot \cdots \boxdot (\widetilde{P}_j, \widetilde{Q}_j, \widetilde{n}_j). \tag{9.169}$$

Similarly, we find $0 \leqslant m \leqslant k$ and $(\widetilde{P}'_m, \widetilde{Q}'_m, \widetilde{n}'_m) \in \mathcal{R}(\widetilde{I}_0)$ such that either $m = k$, $(\widetilde{P}'_m, \widetilde{Q}'_m, \widetilde{n}'_m) = (P_\omega, Q_\omega, n_\omega)$ or $m < k$ and

$$(P_\omega, Q_\omega, n_\omega) \in (\widetilde{P}'_m, \widetilde{Q}'_m, \widetilde{n}'_m) \boxdot \cdots \boxdot (P_k, Q_k, n_k). \tag{9.170}$$

We also have $P_j \subset \widetilde{P}_j$, $Q_m \subset \widetilde{Q}'_m$. Moreover, as (P, Q, n) is free, we must have $j \leqslant m$ and, when $j = m$, we must also have $n_j = n_m > \widetilde{n}_j + \widetilde{n}'_m$.

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 $j = m$, $P_j \subset \widetilde{P}_j$, $Q_j \subset \widetilde{Q}'_j$ and the largest parameter interval $\widehat{I}_1$ for which $(P_j, Q_j, n_j) \in \mathcal{R}(\widehat{I}_1)$ is not the starting interval I_0. We denote by $\widetilde{I}_1$ the parent interval. We rewrite

$$\begin{aligned}
(P^1, Q^1, n^1) &:= (P_j, Q_j, n_j),\\
(P^1_\alpha, Q^1_\alpha, n^1_\alpha) &:= (\widetilde{P}_j, \widetilde{Q}_j, \widetilde{n}_j),\\
(P^1_\omega, Q^1_\omega, n^1_\omega) &:= (\widetilde{P}'_j, \widetilde{Q}'_j, \widetilde{n}'_j),
\end{aligned} \tag{9.171}$$

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 $0 \leqslant j_1 \leqslant m_1 \leqslant k_1$ (with $k_1 > 0$), elements (P^1_i, Q^1_i, n^1_i) for $0 \leqslant i \leqslant k_1$ and also $(P^2_\alpha, Q^2_\alpha, n^2_\alpha)$, $(P^2_\omega, Q^2_\omega, n^2_\omega)$, all in $\mathcal{R}(\widetilde{I}_1)$, such that

$$\begin{aligned}
(P^1, Q^1, n^1) &\in (P^1_0, Q^1_0, n^1_0) \boxdot \cdots \boxdot (P^1_{k_1}, Q^1_{k_1}, n^1_{k_1}),\\
(P^1_\alpha, Q^1_\alpha, n^1_\alpha) &\in (P^1_0, Q^1_0, n^1_0) \boxdot \cdots \boxdot (P^1_{j_1-1}, Q^1_{j_1-1}, n^1_{j_1-1}) \boxdot (P^2_\alpha, Q^2_\alpha, n^2_\alpha),\\
(P^1_\omega, Q^1_\omega, n^1_\omega) &\in (P^2_\omega, Q^2_\omega, n^2_\omega) \boxdot (P^1_{m_1+1}, Q^1_{m_1+1}, n^1_{m_1+1}) \boxdot \cdots\\
&\qquad \boxdot (P^1_{k_1}, Q^1_{k_1}, n^1_{k_1}).
\end{aligned} \tag{9.172}$$

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 $(P^1_{j_1}, Q^1_{j_1}, n^1_{j_1})$ is defined over the starting interval I_0; otherwise we set

$$(P^2, Q^2, n^2) := (P^1_{j_1}, Q^1_{j_1}, n^1_{j_1}), \tag{9.173}$$

and we go on. The sequence of parameter intervals $\widehat{I}_0 \subset \widehat{I}_1 \subset \cdots$ 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.

9.9.4. *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

$$\#Bi_+(P_\alpha, Q_\omega, 0) \leqslant b_+(P_\alpha, 0)b_-(Q_\omega, 0). \tag{9.174}$$

Let $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(\widetilde{I})$, $(P_\omega, Q_\omega, n_\omega) \in \mathcal{C}_-(\widetilde{I})$ and let $(P, Q, n) \in Bi_+(P_\alpha, Q_\omega, 0)$. Then both $(P_\alpha, Q_\alpha, n_\alpha)$, $(P_\omega, Q_\omega, n_\omega) \in \mathcal{R}(I_0)$ and (P, Q, n) is obtained by a simple composition

$$(P, Q, n) = (P_\alpha, Q_\alpha, n_\alpha) * (P', Q', n') * (P_\omega, Q_\omega, n_\omega). \tag{9.175}$$

We have here, by definition of $Bi_+(P_\alpha, Q_\omega)$ and $\mathcal{C}_+(\widetilde{I})$

$$x \leqslant \max_{\widetilde{I}} |P| \leqslant C \max_{\widetilde{I}} |P_\alpha| \max_{\widetilde{I}} |P'| \max_{\widetilde{I}} |P_\omega| \leqslant C|\widetilde{I}|^{1+\tau} \max_{\widetilde{I}} |P'| \max_{\widetilde{I}} |P_\omega|. \tag{9.176}$$

This gives

$$\max_{\widetilde{I}} |P'| \geqslant C^{-1}|\widetilde{I}|^{-(1+\tau)}\Big(\max_{\widetilde{I}} |P_\omega|\Big)^{-1} x. \tag{9.177}$$

From the reminder at the beginning of Section 8.3, we can thus define, as $d_s^0 + C\varepsilon_0 < d_s^* < \rho_s$,

$$b_+(P_\alpha, 0) = \begin{cases} (C|\widetilde{I}|^{1+\tau}x^{-1})^{\rho_s} & \text{if } (P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{R}(I_0),\\ 0 & \text{otherwise,}\end{cases} \tag{9.178}$$

$$b_-(Q_\omega, 0) = \begin{cases} (\max_{\widetilde{I}}|P_\omega|)^{\rho_s} & \text{if } (P_\omega, Q_\omega, n_\omega) \in \mathcal{R}(I_0),\\ 0 & \text{otherwise.}\end{cases} \tag{9.179}$$

Then, (9.174) is satisfied.

9.9.5. *Subclasses of higher depth.* — Let $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{C}_+(\widetilde{I})$, $(P_\omega, Q_\omega, n_\omega) \in \mathcal{C}_-(\widetilde{I})$ 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 Section 9.9.2. We set

$$\begin{aligned} (P^0, Q^0, n^0) &:= (P, Q, n),\\ (P^0_\alpha, Q^0_\alpha, n^0_\alpha) &:= (P_\alpha, Q_\alpha, n_\alpha),\\ (P^0_\omega, Q^0_\omega, n^0_\omega) &:= (P_\omega, Q_\omega, n_\omega).\end{aligned} \tag{9.180}$$

We have

– a strictly increasing sequence of parameter intervals

$$(\mathbf{9.181})\qquad \widehat{I}_0 \subset \widehat{I}_1 \subset \cdots \subset \widehat{I}_{s-1} \subset \widehat{I}_s = I_0$$

with $\widetilde{I} \subset \widehat{I}_0$; we denote by $\widetilde{I}_r$ the parent interval of $\widehat{I}_r$ for $0 \leqslant r < s$;
– a sequence (P^r, Q^r, n^r), $0 \leqslant r \leqslant s$ such that (P^r, Q^r, n^r) belongs to $\mathcal{R}(\widehat{I}_r)$ but not to $\mathcal{R}(\widetilde{I}_r)$ for $r < s$; also $(P^s, Q^s, n^s) \in \mathcal{R}(\widehat{I}_{s-1})$;
– two sequences $(P^r_\alpha, Q^r_\alpha, n^r_\alpha)$, $(P^r_\omega, Q^r_\omega, n^r_\omega)$, $0 \leqslant r \leqslant s$; for each $r < s$, resp. $r = s$, the two elements belong to $\mathcal{R}(\widetilde{I}_r)$, resp. $\mathcal{R}(\widetilde{I}_{s-1})$;
– two sequences (P^r_+, Q^r_+, n^r_+), (P^r_-, Q^r_-, n^r_-), $0 < r \leqslant s$; for each r, the two elements belong to $\mathcal{R}(\widehat{I}_{r-1})$.

These data are related by the following properties: for each $0 < r \leqslant s$, we have

$$(\mathbf{9.182})\qquad (P^{r-1}, Q^{r-1}, n^{r-1}) \in (P^r_-, Q^r_-, n^r_-) \boxdot (P^r, Q^r, n^r) \boxdot (P^r_+, Q^r_+, n^r_+),$$

$$(\mathbf{9.183})\qquad (P^{r-1}_\alpha, Q^{r-1}_\alpha, n^{r-1}_\alpha) \in (P^r_-, Q^r_-, n^r_-) \boxdot (P^r_\alpha, Q^r_\alpha, n^r_\alpha),$$

$$(\mathbf{9.184})\qquad (P^{r-1}_\omega, Q^{r-1}_\omega, n^{r-1}_\omega) \in (P^r_\omega, Q^r_\omega, n^r_\omega) \boxdot (P^r_+, Q^r_+, n^r_+).$$

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 $\mathcal{R}(\widetilde{I}_{s-1})$; then, by the structure theorem, there exists an integer $h > 0$, elements $(P^s_0, Q^s_0, n^s_0) \cdots (P^s_h, Q^s_h, n^s_h)$ in $\mathcal{R}(\widetilde{I}_{s-1})$ with

$$(\mathbf{9.185})\qquad (P^s, Q^s, n^s) \in (P^s_0, Q^s_0, n^s_0) \boxdot \cdots \boxdot (P^s_h, Q^s_h, n^s_h)$$

and also

$$(\mathbf{9.186})\qquad P^s_0 \subset P^s_\alpha, \qquad Q^s_h \subset Q^s_\omega.$$

(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 (9.182) through (9.184) take place in $\mathcal{R}(\widehat{I}_{r-1})$ but not in $\mathcal{R}(\widetilde{I}_{r-1})$; in (9.185), they take place in $\mathcal{R}(\widehat{I}_{s-1})$ but not in $\mathcal{R}(\widetilde{I}_{s-1})$.

A subclass $\mathrm{B}i_+(P_\alpha, Q_\omega, \ell)$, 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 $\widehat{I}_0 \subset \cdots \subset \widehat{I}_s = I_0$;
– the integer $h \geqslant 0$;
– when $h > 1$, for each $0 < i < h$, the largest negative integral power of 2, denoted by x_i, such that $\max_{\widetilde{I}} |P^s_i| \geqslant x_i$;
– when $h > 0$, the largest negative integral powers of 2, denoted by x_0, x_h, such that $\max_{\widetilde{I}} |P_-| \geqslant x_0$, $\max_{\widetilde{I}} |P_+| \geqslant x_h$; here, the elements (P_-, Q_-, n_-), (P_+, Q_+, n_+) are determined by $P \subset P_-$, $Q \subset Q_+$ and

$$(\mathrm{P}_-, \mathrm{Q}_-, n_-) \in (\mathrm{P}^1_-, \mathrm{Q}^1_-, n^1_-) \,\square \cdots \square\, (\mathrm{P}^s_-, \mathrm{Q}^s_-, n^s_-) \,\square\, (\mathrm{P}^s_0, \mathrm{Q}^s_0, n^s_0), \tag{9.187}$$

$$(\mathrm{P}_+, \mathrm{Q}_+, n_+) \in (\mathrm{P}^s_h, \mathrm{Q}^s_h, n^s_h) \,\square\, (\mathrm{P}^s_+, \mathrm{Q}^s_+, n^s_+) \,\square \cdots \square\, (\mathrm{P}^1_+, \mathrm{Q}^1_+, n^1_+). \tag{9.188}$$

One has by construction $\mathrm{P}_- \subset \mathrm{P}_\alpha$, $\mathrm{Q}_+ \subset \mathrm{Q}_\omega$.

Thus, we group together in a subclass $\mathrm{B}i_+(\mathrm{P}_\alpha, \mathrm{Q}_\omega, \ell)$ the elements of $\mathrm{B}i_+(\mathrm{P}_\alpha, \mathrm{Q}_\omega)$ who share the same data; the elements of L, distinct from $\lozenge$, 0, are the sets of data for which at least one subclass $\mathrm{B}i_+(\mathrm{P}_\alpha, \mathrm{Q}_\omega, \ell)$ is non-empty, for some $(\mathrm{P}_\alpha, \mathrm{Q}_\alpha, n_\alpha)$ in $\mathcal{C}_+(\widetilde{\mathrm{I}})$, $(\mathrm{P}_\omega, \mathrm{Q}_\omega, n_\omega)$ in $\mathcal{C}_-(\widetilde{\mathrm{I}})$.

The definition of the set L is now complete.

9.9.6. *Sizes of subclasses of higher depth.* — The context and notations are the same as above. We want to define $b_+(\mathrm{P}_\alpha, \ell)$ and $b_-(\mathrm{Q}_\omega, \ell)$ in order to satisfy (9.124) in Section 9.8.

We first observe that $(\mathrm{P}_\alpha, \mathrm{Q}_\alpha, n_\alpha)$ determines $(\mathrm{P}^1_-, \mathrm{Q}^1_-, n^1_-), \dots, (\mathrm{P}^s_-, \mathrm{Q}^s_-, n^s_-)$, $(\mathrm{P}^s_\alpha, \mathrm{Q}^s_\alpha, n^s_\alpha)$ and the result of parabolic compositions between these elements. Similarly, $(\mathrm{P}_\omega, \mathrm{Q}_\omega, n_\omega)$ determines $(\mathrm{P}^1_+, \mathrm{Q}^1_+, n^1_+), \dots, (\mathrm{P}^s_+, \mathrm{Q}^s_+, n^s_+)$, $(\mathrm{P}^s_\omega, \mathrm{Q}^s_\omega, n^s_\omega)$ and the result of parabolic compositions between these elements. Therefore, the only "freedom" for the element $(\mathrm{P}, \mathrm{Q}, n)$ in the subclass $\mathrm{B}i_+(\mathrm{P}_\alpha, \mathrm{Q}_\omega, \ell)$ is through $(\mathrm{P}^s, \mathrm{Q}^s, n^s)$, and this freedom is constrained by the relations $\mathrm{P}^s \subset \mathrm{P}^s_\alpha$, $\mathrm{Q}^s \subset \mathrm{Q}^s_\omega$.

Consider first a subclass with $h = 0$, i.e., $(\mathrm{P}^s, \mathrm{Q}^s, n^s) \in \mathcal{R}(\mathrm{I}_0)$. The widths of the strips are related as follows: for every $t \in \widetilde{\mathrm{I}}$, we have

$$\mathrm{C}^{-1} \frac{|\mathrm{P}^s|}{|\mathrm{P}^s_\alpha||\mathrm{P}^s_\omega|} \leqslant \frac{|\mathrm{P}|}{|\mathrm{P}_\alpha||\mathrm{P}_\omega|} \leqslant \mathrm{C} \frac{|\mathrm{P}^s|}{|\mathrm{P}^s_\alpha||\mathrm{P}^s_\omega|}. \tag{9.189}$$

This allows us to take, as in the case of depth 0,

$$b_+(\mathrm{P}_\alpha, \ell) = \begin{cases} (\mathrm{C}|\widetilde{\mathrm{I}}|^{1+\tau} x^{-1})^{\rho_s} & \text{if } (\mathrm{P}^s_\alpha, \mathrm{Q}^s_\alpha, n^s_\alpha) \in \mathcal{R}(\mathrm{I}_0), \\ 0 & \text{otherwise,} \end{cases} \tag{9.190}$$

$$b_-(\mathrm{Q}_\omega, \ell) = \begin{cases} \max_{\widetilde{\mathrm{I}}} |\mathrm{P}_\omega|^{\rho_s} & \text{if } (\mathrm{P}^s_\omega, \mathrm{Q}^s_\omega, n^s_\omega) \in \mathcal{R}(\mathrm{I}_0), \\ 0 & \text{otherwise.} \end{cases} \tag{9.191}$$

Consider now a subclass with $h > 0$, i.e., case (a) in Section 9.9.4 above.

By Lemma 10 (Section 9.1), for $h > 1$ and $0 < i < h$, P^s_i and Q^s_i are thin $\widetilde{\mathrm{I}}_{s-1}$-critical, hence element $(\mathrm{P}^s_i, \mathrm{Q}^s_i, n^s_i)$ belongs to $\mathrm{B}i_+(\widetilde{\mathrm{I}}_{s-1}, \widetilde{\mathrm{I}}_{s-1}, \widetilde{\mathrm{I}}_{s-1}; x_i)$. In the parabolic composition

$$(\mathrm{P}, \mathrm{Q}, n) \in (\mathrm{P}_-, \mathrm{Q}_-, n_-) \,\square\, (\mathrm{P}^s_1, \mathrm{Q}^s_1, n^s_1) \,\square \cdots \square\, (\mathrm{P}^s_{h-1}, \mathrm{Q}^s_{h-1}, n^s_{h-1}) \,\square\, (\mathrm{P}_+, \mathrm{Q}_+, n_+),$$

we have $\delta(Q_-, P_1^s) > |\widehat{I}_{s-1}|, \dots, \delta(Q_{h-1}^s, P_+) > |\widehat{I}_{s-1}|$ for all $t \in \widetilde{I}$ by Lemma 3 (Section 6.6.3), hence, using (3.27)

$$\max_{\widetilde{I}} |P| \leqslant C^h |\widehat{I}_{s-1}|^{-h/2} \max_{\widetilde{I}} |P_-| \max_{\widetilde{I}} |P_1^s| \cdots \max_{\widetilde{I}} |P_+|.$$

This gives, by definition of the x_i

$$x \leqslant \left(C |\widehat{I}_{s-1}|^{-\frac{1}{2}} \right)^h x_0 x_1 \cdots x_h. \tag{9.192}$$

Thus, the data of every subclass must satisfy (9.192). Assuming that (9.192) holds, we set $b_+(P_\alpha, \ell) = 0$ if $(P_\alpha, Q_\alpha, n_\alpha) \notin \mathcal{R}(\widehat{I}_0)$. When $(P_\alpha, Q_\alpha, n_\alpha) \in \mathcal{R}(\widehat{I}_0)$, we set

$$b_+(P_\alpha, \ell) = 2^h \Big(\prod_{0<i<h} \# Bi_+(\widetilde{I}_{s-1}, \widetilde{I}_{s-1}, \widetilde{I}_{s-1}; x_i) \Big) \# Bi_+(P_\alpha, \widetilde{I}_{s-1}; x_0). \tag{9.193}$$

Here, $Bi_+(P_\alpha, \widetilde{I}_{s-1}, x_0)$ is by definition the set of elements (P_-, Q_-, n_-) in $\mathcal{R}(\widetilde{I})$ such that $P_- \subset P_\alpha$, Q_- is thin $\widetilde{I}_{s-1}$-critical and $\max_{\widetilde{I}} |P_-| \geqslant x_0$.

Similarly, when (9.192) holds, we set $b_-(Q_\omega, \ell) = 0$ if $(P_\omega, Q_\omega, n_\omega) \notin \mathcal{R}(\widehat{I}_0)$. When $(P_\omega, Q_\omega, n_\omega) \in \mathcal{R}(\widehat{I}_0)$, we set

$$b_-(Q_\omega, \ell) = \# Bi_+(\widetilde{I}_{s-1}, Q_\omega; x_h), \tag{9.194}$$

where now $Bi_+(\widetilde{I}_{s-1}, Q_\omega; x_h)$ is the set of elements (P_+, Q_+, n_+) in $\mathcal{R}(\widetilde{I})$ such that $Q_+ \subset Q_\omega$, P_+ is thin $\widetilde{I}_{s-1}$-critical and $\max_{\widetilde{I}} |P_+| \geqslant x_h$.

The factor 2^h in (9.193) 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 (9.124) is satisfied.

9.10. *The size of the index set* L. — It is not difficult from (9.192) to see that the index set L is finite, but we need an explicit bound on its cardinality (cf. (9.134)).

We assume that, with C large enough,

$$\rho_0 > d_s^* + C\tau^{-2} \frac{\log\log \varepsilon_0^{-1}}{\log \varepsilon_0^{-1}}, \tag{9.195}$$

where $\frac{\log\log \varepsilon_0^{-1}}{\log \varepsilon_0^{-1}} \ll \tau^2$ when ε_0 is small enough.

Proposition **45.** — *The index set* L *satisfies*

$$\# L \leqslant |\widetilde{I}|^{-\tau^2}.$$

Proof. — It is sufficient to count the subclasses of strictly positive depth (there are only two subclasses besides these ones).

In a first step, we fix the depth $s > 0$ and the sequence of intervals $\widehat{I}_0 \subset \cdots \subset \widehat{I}_{s-1} \subset \widehat{I}_s = I_0$. There is one subclass with $h = 0$ and we will estimate the number of subclasses with $h > 0$ (the integer h itself is *not* fixed). This number is bounded by the number of $(h+1)$-tuples $(x_0, \ldots, x_h)$ (of negative integral powers of 2) which satisfy (9.192).

By Proposition 14 (Section 6.6.3), we have

$$x_i < |\widetilde{I}_{s-1}|^{\beta} \quad \text{for } 0 < i < h, \tag{9.196}$$

$$x_h < |\widetilde{I}_{s-1}|. \tag{9.197}$$

As $P_- \subset P_\alpha$, we also have, for a non-empty subclass

$$x_0 < |\widetilde{I}|^{1+\tau}. \tag{9.198}$$

We rewrite (9.192) as

$$\frac{x_0}{|\widetilde{I}|^{1+\tau}} \left(\prod_{0<i<h} \frac{x_i}{|\widetilde{I}_{s-1}|^{\beta}} \right) \frac{x_h}{|\widetilde{I}_{s-1}|} \geqslant \frac{x}{|\widetilde{I}|^{1+\tau} |\widetilde{I}_{s-1}|^{\beta(h-1)+1}} \left(C^{-1} |\widehat{I}_{s-1}|^{\frac{1}{2}} \right)^h. \tag{9.199}$$

Using $\beta > 1$, and taking base-two logarithms, it is sufficient to bound the number of non-negative integral solutions of

$$n_0 + \cdots + n_h \leqslant A_0 - A_1 h, \tag{9.200}$$

with

$$A_0 = \log_2(|\widetilde{I}| x^{-1}), \tag{9.201}$$

$$A_1 = \frac{1}{3} \log_2 |\widehat{I}_{s-1}|^{-1}. \tag{9.202}$$

We have $x \leqslant x^*_{cr} \leqslant \widetilde{x}_{cr} = \varepsilon_0 |P_u| (\frac{|\widetilde{I}|}{\varepsilon_0})^{\frac{\sigma_0}{\rho_0 - \rho_1}} \ll |\widetilde{I}|$, because $\sigma_0 > \rho_0 - \rho_1$ and $|P_u| \ll 1$. Therefore A_0 is large; it is obvious that A_1 is also 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 (9.200) (including those with $h = 0$) is then the coefficient of z^{A_0} in the power series for

$$\chi(z) := \sum_{h \geqslant 0} z^{A_1 h} (1-z)^{-h-2} = (1-z)^{-1} (1 - z - z^{A_1})^{-1}. \tag{9.203}$$

We estimate this coefficient by a Cauchy integral on the circle $|z| = 1 - 2A_1^{-1} \log A_1$. On this circle, we have

$$|z^{A_1}| < A_1^{-1}, \tag{9.204}$$

$$|\chi(z)| < \frac{1}{2}A_1^2(\log A_1)^{-2}. \tag{9.205}$$

The number of solutions of (9.200) is therefore not greater than

$$A_1^2(\log A_1)^{-2}(1-2A_1^{-1}\log A_1)^{-A_0}. \tag{9.206}$$

In view of (9.202), this quantity is smaller than

$$(\log|\widehat{I}_{s-1}|)^2\exp\left(CA_0\frac{\log|\log|\widehat{I}_{s-1}||}{|\log|\widehat{I}_{s-1}||}\right). \tag{9.207}$$

This is a bound for the number of subclasses with fixed depth s and fixed sequence $\widehat{I}_0 \subset \cdots \subset \widehat{I}_{s-1}$. We have now to sum over these remaining data. Observe that (9.207) depends on $|\widehat{I}_{s-1}|$, *not* on the depth s and the intervals $\widehat{I}_r$, $0 \leqslant r < s-1$.

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 $\widetilde{I} \subset I^* \subset \widehat{I}$, $I^* \neq \widehat{I}$. 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 (with $s > 0$) is bounded by

$$\sum_{\widehat{I}} 2^{S(\widehat{I})}(\log|\widehat{I}|)^2\exp\left(CA_0\frac{\log|\log|\widehat{I}||}{|\log|\widehat{I}||}\right). \tag{9.208}$$

We have here

$$\frac{\log|\log|\widehat{I}||}{|\log|\widehat{I}||} \leqslant \frac{\log\log\varepsilon_0^{-1}}{\log\varepsilon_0^{-1}}, \tag{9.209}$$

$$|\log|\widetilde{I}|| = (1+\tau)^{S(\widehat{I})}|\log|\widehat{I}||, \tag{9.210}$$

$$S(\widehat{I}) \leqslant 2\tau^{-1}\log_2\left(\frac{\log|\widetilde{I}|^{-1}}{\log\varepsilon_0^{-1}}\right) =: S_{\max}. \tag{9.211}$$

The sum (9.208) is thus bounded by

$$C2^{S_{\max}}(\log\varepsilon_0^{-1})^2\exp\left(CA_0\frac{\log\log\varepsilon_0^{-1}}{\log\varepsilon_0^{-1}}\right)$$

$$\leqslant (\log|\widetilde{I}|^{-1})^{2\tau^{-1}}\exp\left(CA_0\frac{\log\log\varepsilon_0^{-1}}{\log\varepsilon_0^{-1}}\right). \tag{9.212}$$

As $x \geqslant x_{\min} := |\widetilde{I}|^{C(\rho_0-d_s^*)^{-1}}$, we have

$$A_0 \leqslant C(\rho_0-d_s^*)^{-1}\log|\widetilde{I}|^{-1}. \tag{9.213}$$

If the constant in (9.195) is large enough, we obtain

$$\text{(9.214)} \qquad CA_0 \frac{\log\log \varepsilon_0^{-1}}{\log \varepsilon_0^{-1}} < \frac{1}{2}\tau^2 \log|\widetilde{I}|^{-1}.$$

Introducing this in (9.212) concludes the proof of the proposition. □

9.11. *The size of* $\widehat{B}$.

9.11.1. According to the roadmap exposed in Section 9.7.6, we have now to estimate the quantity set in Section 9.7.6

$$\text{(9.140)} \qquad \widehat{B} = \sum_{L} b_+(\ell) b_-(\ell)$$

with

$$\text{(9.136)} \qquad b_+(\ell) = \sum_{\mathcal{C}_+(\widetilde{I})} b_+(P_\alpha, \ell),$$

$$\text{(9.137)} \qquad b_-(\ell) = \sum_{\mathcal{C}_-(\widetilde{I})} b_-(Q_\omega, \ell).$$

Consider first the bound elements. In view of (9.173), we have:

$$\text{(9.215)} \qquad \begin{aligned} b_+(\Diamond) &= \#\mathcal{C}_+(\widetilde{I})(\rho_0 - d_s^*)^{-1} \log|\widetilde{I}|^{-1}, \\ b_-(\Diamond) &= \#\mathcal{C}_-(\widetilde{I})(\rho_0 - d_s^*)^{-1} \log|\widetilde{I}|^{-1}. \end{aligned}$$

Consider next the class of depth 0, and also the classes of higher depth with $h = 0$: in view of (9.178)–(9.179) and (9.190)–(9.191), we have in these cases

$$\text{(9.216)} \qquad b_+(\ell) \leqslant (C|\widetilde{I}|^{1+\tau} x^{-1})^{\rho_s} \#\mathcal{C}_+(\widetilde{I}),$$

$$\text{(9.217)} \qquad b_-(\ell) \leqslant \sum_{\mathcal{C}_-(\widetilde{I})} \left(\max_{\widetilde{I}} |P_\omega|\right)^{\rho_s}.$$

Also, the number of such classes, according to the discussion in the proof of Proposition 45 is not larger than

$$\text{(9.218)} \qquad 2^{S_{\max}} \leqslant \left(\frac{\log|\widetilde{I}|^{-1}}{\log \varepsilon_0^{-1}}\right)^{2\tau^{-1}}.$$

9.11.2. The remaining subclasses are more complicated! Formulas (9.193), (9.194) suggest an induction. We thus assume that $(\mathrm{SR3})_s$ is satisfied for all parameter intervals containing $\widetilde{\mathrm{I}}$. We have, for a class of depth $s > 0$ with $h > 0$:

$$b_+(\ell) \leqslant 2^h \left(\prod_{0<i<h} (\#\mathrm{B}i_+(\widetilde{\mathrm{I}}_{s-1}, \widetilde{\mathrm{I}}_{s-1}, \widetilde{\mathrm{I}}_{s-1}; x_i)) \right) \#\mathrm{B}i_+(\widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}_{s-1}; x_0), \tag{9.219}$$

$$b_-(\ell) \leqslant \#\mathrm{B}i_+(\widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}_{s-1}, \widetilde{\mathrm{I}}; x_h). \tag{9.220}$$

We recall from (9.199) that

$$x \leqslant \left(\mathrm{C}|\widehat{\mathrm{I}}_{s-1}|^{-\frac{1}{2}} \right)^h x_0 x_1 \cdots x_h.$$

Observe that, from (9.16), the critical value $\widehat{x}_{cr}$ in each of the $\mathrm{B}i_+$ sets above is the same and equal to

$$\widehat{x}_{cr} := \varepsilon_0 |\mathrm{P}_u| \left(\frac{|\widetilde{\mathrm{I}}_{s-1}|}{\varepsilon_0} \right)^{\frac{\sigma_0}{\rho_0 - \rho_1}}. \tag{9.221}$$

As in Section 9.8, we separate the subclasses into two parts: those for which every x_i is above the critical value $\widehat{x}_{cr}$ and the others.

9.11.3. *Subclasses with all x_i large.* — In this case, we have from $(\mathrm{SR3})_s$

$$\#\mathrm{B}i_+(\widetilde{\mathrm{I}}_{s-1}, \widetilde{\mathrm{I}}_{s-1}, \widetilde{\mathrm{I}}_{s-1}; x_i) \leqslant \varepsilon_0^{-\mathrm{A}\tau} \left(\frac{x_i}{\varepsilon_0 |\mathrm{P}_u|} \right)^{-\rho_1} \left(\frac{|\widetilde{\mathrm{I}}_{s-1}|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1}, \quad \text{for } 0 < i < h, \tag{9.222}$$

$$\#\mathrm{B}i_+(\widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}_{s-1}; x_0) \leqslant \varepsilon_0^{-\mathrm{A}\tau} \left(\frac{x_0}{\varepsilon_0 |\mathrm{P}_u|} \right)^{-\rho_1} \left(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1}, \tag{9.223}$$

$$\#\mathrm{B}i_+(\widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}_{s-1}, \widetilde{\mathrm{I}}; x_h) \leqslant \varepsilon_0^{-\mathrm{A}\tau} \left(\frac{x_h}{\varepsilon_0 |\mathrm{P}_u|} \right)^{-\rho_1} \left(\frac{|\widetilde{\mathrm{I}}_{s-1}|}{\varepsilon_0} \right)^{\sigma_1} \left(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0} \right)^{\sigma_0}. \tag{9.224}$$

Multiplying these inequalities, we obtain, taking (9.192) into account

$$b_+(\ell) b_-(\ell) \leqslant \mathrm{A}_2^h \mathrm{A}_3, \tag{9.225}$$

with

$$\mathrm{A}_2 = 2\varepsilon_0^{-2\mathrm{A}\tau} \left(\frac{|\widetilde{\mathrm{I}}_{s-1}|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1} \left(\mathrm{C}\varepsilon_0 |\mathrm{P}_u| |\widehat{\mathrm{I}}_{s-1}|^{-\frac{1}{2}} \right)^{\rho_1}, \tag{9.226}$$

$$\mathrm{A}_3 = \mathrm{C}^{\rho_1} \left(\frac{x}{\varepsilon_0 |\mathrm{P}_u|} \right)^{-\rho_1} \left(\frac{|\widetilde{\mathrm{I}}|}{\varepsilon_0} \right)^{\sigma_0 + \sigma_1}. \tag{9.227}$$

9.11.4. *Subclasses with some x_i small.*

Lemma **13.** — *Let* $I \subset \widetilde{I}$ *be a candidate interval. Let* $(P, Q, n) \in \mathcal{R}(I)$. *If* P *and* Q *are thin* I_0*-critical, then we have* $\max_I |P| \leqslant C\varepsilon_0 |P_u|$ *and* $\max_I |Q| \leqslant C\varepsilon_0 |Q_s|$.

Proof. — We prove the first statement by induction on the level of the largest interval I_1 such that $(P, Q, n) \in \mathcal{R}(I_1)$. If $I_1 = I_0$, the statement follows from the fact that (P, Q, n) is the simple composition of (P_s, Q_s, n_s), some $(P', Q', n') \in \mathcal{R}(I_0)$ and (P_u, Q_u, n_u) (see Section 9.4). If $I_1 \neq I_0$ and $\widetilde{I}_1$ is the parent interval, consider the thinnest $\widetilde{I}_1$-defined rectangle P_0 containing P. As P is thin I_0-critical, P_0 is also thin I_0-critical. By Lemma 10 in Section 9.1, Q_0 is thin I_0-critical. By the induction hypothesis, we have $\max_{\widetilde{I}_1} |P_0| \leqslant C\varepsilon_0 |P_u|$. As $P \subset P_0$, we are able to conclude.

The proof of the other statement is symmetric. □

For $0 < i < h$, the cardinality of $Bi_+(\widetilde{I}_{s-1}, \widetilde{I}_{s-1}, \widetilde{I}_{s-1}; x_i)$ is controlled by (9.229) if $x_i \geqslant \widehat{x}_{cr}$ and by

$$\#Bi_+(\widetilde{I}_{s-1}, \widetilde{I}_{s-1}, \widetilde{I}_{s-1}; x_i) \leqslant \varepsilon_0^{-A\tau} \left(\frac{x_i}{\varepsilon_0 |P_u|}\right)^{-\rho_0} \left(\frac{|\widetilde{I}_{s-1}|}{\varepsilon_0}\right)^{2\sigma_0+\sigma_1}, \tag{9.228}$$

if $x_i \leqslant \widehat{x}_{cr}$. Similarly, we have (9.230) if $x_0 \geqslant \widehat{x}_{cr}$ and

$$\#Bi_+(\widetilde{I}, \widetilde{I}, \widetilde{I}_{s-1}; x_0) \leqslant \varepsilon_0^{-A\tau} \left(\frac{x_0}{\varepsilon_0 |P_u|}\right)^{-\rho_0} \left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{\sigma_0+\sigma_1} \left(\frac{|\widetilde{I}_{s-1}|}{\varepsilon_0}\right)^{\sigma_0}, \tag{9.229}$$

if $x_0 \leqslant \widehat{x}_{cr}$; we have (9.231) if $x_h \geqslant \widehat{x}_{cr}$ and

$$\#Bi_+(\widetilde{I}, \widetilde{I}_{s-1}, \widetilde{I}; x_h) \leqslant \varepsilon_0^{-A\tau} \left(\frac{x_h}{\varepsilon_0 |P_u|}\right)^{-\rho_0} \left(\frac{|\widetilde{I}_{s-1}|}{\varepsilon_0}\right)^{\sigma_0+\sigma_1} \left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{\sigma_0} \tag{9.230}$$

if $x_0 \leqslant \widehat{x}_{cr}$.

By Lemma 13, all the terms $\frac{x_j}{\varepsilon_0 |P_u|}$ are bounded by C. As $\rho_0 > \rho_1$, we can when necessary replace the exponent ρ_1 (for those x_j which are larger than $\widehat{x}_{cr}$) by ρ_0. On the other hand comparing the σ exponents, we see that we always have an additional factor $(\frac{|\widetilde{I}_{s-1}|}{\varepsilon_0})^{\sigma_0}$ when $x_j \leqslant \widehat{x}_{cr}$ compared with $x_j \geqslant \widehat{x}_{cr}$ (including the cases $j = 0, h$). As at least one x_j is $\leqslant \widehat{x}_{cr}$, multiplying the estimates together gives

$$b_+(\ell) b_-(\ell) \leqslant \widetilde{A}_2^h \widetilde{A}_3, \tag{9.231}$$

with

$$\widetilde{A}_2 = 2\varepsilon_0^{-2A\tau} \left(\frac{|\widetilde{I}_{s-1}|}{\varepsilon_0}\right)^{\sigma_0+\sigma_1} \left(C\varepsilon_0 |P_u| |\widehat{I}_{s-1}|^{-\frac{1}{2}}\right)^{\rho_0}, \tag{9.232}$$

$$\widetilde{A}_3 = C^{\rho_0} \left(\frac{x}{\varepsilon_0 |P_u|}\right)^{-\rho_0} \left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{2\sigma_0+\sigma_1}. \tag{9.233}$$

9.11.5. *Partial sums of the $b_+(\ell)b_-(\ell)$.* — We observe that in both (9.225) and (9.231), our estimate for $b_+(\ell)b_-(\ell)$ depends on the class ℓ only through $\widehat{\mathrm{I}}_{s-1}$ and h.

We first sum over subclasses with a fixed depth s and fixed sequence $\widehat{\mathrm{I}}_0 \subset \cdots \subset \widehat{\mathrm{I}}_{s-1}$, using the same method of generating series as in the proof of Proposition 45.

To deal with the two cases above at the same time, we observe that $\sigma_0 + \sigma_1 - \frac{1}{2}\rho_0(1+\tau)$ is close to $1 - d^0_u - \frac{1}{2}d^0_s$ when $\tau \gg \eta \gg \varepsilon_0$ are small, with $1 - d^0_u - \frac{1}{2}d^0_s > 0$ from (H4). A fortiori $\sigma_0 + \sigma_1 - \frac{1}{2}\rho_1(1+\tau) > 0$ is positive and bounded away from 0. Thus, A_2 and $\widetilde{\mathrm{A}}_2$ are larger when $\widetilde{\mathrm{I}}_{s-1}$ is larger; the largest case is $\widetilde{\mathrm{I}}_{s-1} = \mathrm{I}_0$, which gives

$$\max(\mathrm{A}_2, \widetilde{\mathrm{A}}_2) \leqslant \widehat{\mathrm{A}}_2 := 2\varepsilon_0^{-2\mathrm{A}\tau}\Big(\mathrm{C}\varepsilon_0^{\frac{1}{2}(1-\tau)}|\mathrm{P}_u|\Big)^{\rho_1}. \tag{9.234}$$

We have $\max(\mathrm{A}_3, \widetilde{\mathrm{A}}_3) \leqslant \mathrm{C}\widetilde{\mathrm{B}}$, where $\varepsilon_0^{-\mathrm{A}\tau}\widetilde{\mathrm{B}}$ is the bound from $(\mathrm{SR3})_s$ for the cardinality of $\mathrm{B}i_+(\widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}, \widetilde{\mathrm{I}}; x)$.

We set

$$\chi_1(z) = \sum_{h>0} \widehat{\mathrm{A}}_2^h z^{\mathrm{A}_1 h}(1-z)^{-h-2}$$
$$= \widehat{\mathrm{A}}_2 z^{\mathrm{A}_1}(1-z)^{-2}(1 - z - \widehat{\mathrm{A}}_2 z^{\mathrm{A}_1})^{-1}. \tag{9.235}$$

The partial sum of $b_+(\ell)b_-(\ell)$ is thus not larger than $\mathrm{C}\widetilde{\mathrm{B}}$ times the coefficient of z^{A_0} in the power series for $\chi_1(z)$. Recall that A_0, A_1 were defined in (9.201), (9.202).

We estimate this coefficient by Cauchy integration on the circle $\{|z| = 1 - \mathrm{A}_0^{-1} - \widehat{\mathrm{A}}_2\}$, on which we have

$$|1-z|^{-2} \leqslant (\widehat{\mathrm{A}}_2 + \mathrm{A}_0^{-1})^{-2} \leqslant \mathrm{A}_0^2, \tag{9.236}$$

$$|1 - z - \widehat{\mathrm{A}}_2 z^{\mathrm{A}_1}|^{-1} \leqslant \mathrm{A}_0, \tag{9.237}$$

$$|\chi_1(z)| \leqslant \widehat{\mathrm{A}}_2\mathrm{A}_0^3, \tag{9.238}$$

$$|z^{-\mathrm{A}_0}| \leqslant \mathrm{C}(1+\widehat{\mathrm{A}}_2)^{\mathrm{A}_0}. \tag{9.239}$$

The partial sum of $b_+(\ell)b_-(\ell)$ is therefore dominated by

$$\mathrm{C}(1+\widehat{\mathrm{A}}_2)^{\mathrm{A}_0}\widehat{\mathrm{A}}_2\mathrm{A}_0^3\widetilde{\mathrm{B}}. \tag{9.240}$$

9.11.6. *Sum of the $b_+(\ell)b_-(\ell)$ over all free subclasses with $h > 0$.* — We now have to sum over sequences $\widehat{\mathrm{I}}_0 \subset \cdots \subset \widehat{\mathrm{I}}_{s-1}$ and depth s; but (9.240) is independent of these data and the same remarks as in the proof of Proposition 45 apply. So, we finally obtain for the sum of $b_+(\ell)b_-(\ell)$ over subclasses with $s > 0$ and $h > 0$, a bound by

$$\mathrm{C}(1+\widehat{\mathrm{A}}_2)^{\mathrm{A}_0}\widehat{\mathrm{A}}_2\mathrm{A}_0^3 2^{\mathrm{S}_{\max}}\widetilde{\mathrm{B}} \tag{9.241}$$

with $\mathrm{S}_{\max} := 2\tau^{-1}\log_2(\frac{\log|\widetilde{\mathrm{I}}|^{-1}}{\log\varepsilon_0^{-1}})$ (cf. (9.211)).

Recall that (cf. (9.213))

$$(\mathbf{9.242}) \qquad A_0 \leqslant C(\rho_0 - d_s^*)^{-1} \log |\widetilde{I}|^{-1}.$$

As $|P_u|$ is of the order of $\varepsilon_0^{\omega_u}$, we have

$$(\mathbf{9.243}) \qquad \widehat{A}_2 \leqslant \varepsilon_0^{\rho_1/2}.$$

This gives

$$(\mathbf{9.244}) \qquad (1+\widehat{A}_2)^{A_0} \leqslant |\widetilde{I}|^{-C\varepsilon_0^{\rho_1/2}(\rho_0 - d_s^*)^{-1}}.$$

From (9.195), we have

$$(\mathbf{9.245}) \qquad (\rho_0 - d_s^*)^{-1} \ll \tau^2 \log \varepsilon_0^{-1}.$$

The bound that we finally obtain (from (9.241)) for the sum of the $b_+(\ell)b_-(\ell)$ over subclasses with $s > 0$ and $h > 0$ is therefore

$$(\mathbf{9.246}) \qquad \varepsilon_0^{\rho_1/3} |\widetilde{I}|^{-\varepsilon_0^{\rho_1/3}} (\log |\widetilde{I}|^{-1})^{3\tau^{-1}} \widetilde{B}.$$

9.11.7. We summarize the calculations in this subsection in

Proposition **46.** — *The quantity* $\widehat{B} = \sum_L b_+(\ell) b_-(\ell)$ *is bounded by* $\widehat{B}_1 + \widehat{B}_2 + \widehat{B}_3$, *with*

$$\widehat{B}_1 = (\#\mathcal{C}_+(\widetilde{I}))(\#\mathcal{C}_-(\widetilde{I}))(\log |\widetilde{I}|^{-1})^2 (\rho_0 - d_s^*)^{-2},$$

$$\widehat{B}_2 = \left(\frac{\log |\widetilde{I}|^{-1}}{\log \varepsilon_0^{-1}}\right)^{2\tau^{-1}} (C|\widetilde{I}|^{1+\tau} x^{-1})^{\rho_s} (\#\mathcal{C}_+(\widetilde{I})) \sum_{\mathcal{C}_-(\widetilde{I})} \Big(\max_{\widetilde{I}} |P_\omega|\Big)^{\rho_s},$$

$$\widehat{B}_3 = \varepsilon_0^{\rho_1/3} |\widetilde{I}|^{-\varepsilon_0^{\rho_1/3}} (\log |\widetilde{I}|^{-1})^{3\tau^{-1}} \widetilde{B}.$$

9.12. *End of the induction step for* $(\mathrm{SR3})_s$.

9.12.1. In order to complete the induction step for $(\mathrm{SR3})_s$, it is sufficient, in view of (9.136) (Section 9.7.6), to show that

$$(\mathbf{9.247}) \qquad \widehat{B} |\widetilde{I}|^{\tau(2 - d_s^+ - d_u^+ - 6\tau)} \leqslant \varepsilon_0^{-A\tau} B$$

where B is the bound in $(\mathrm{SR3})_s$.

We will bound each of the three expressions $\widehat{B}_i \varepsilon_0^{A\tau} |\widetilde{I}|^{\tau(2-d_s^+ - d_u^+ - 6\tau)}$, $i = 1, 2, 3$, with $\widehat{B}_i$ defined in Proposition 46.

9.12.2. First consider $\widehat{B}_3$. For $x \leqslant \widetilde{x}_{cr}$, we have

$$\widetilde{B} = \widetilde{B}_0 = |\widetilde{I}|^{-\tau(2\sigma_0+\sigma_1)} B_0 \leqslant |\widetilde{I}|^{-\tau(2\sigma_0+\sigma_1)} B, \tag{9.248}$$

and, therefore,

$$\widehat{B}_3 |\widetilde{I}|^{\tau(2-d_s^+-d_u^+-6\tau)} \leqslant \varepsilon_0^{\rho_1/3} (\log |\widetilde{I}|^{-1})^{3\tau^{-1}} |\widetilde{I}|^{\omega} B, \tag{9.249}$$

with $\omega = \tau(2 - d_s^+ - d_u^+ - 6\tau - 2\sigma_0 - \sigma_1) - \varepsilon_0^{\rho_1/3}$.

We choose the exponents σ_0, σ_1 in order to have

$$\omega > \tau^2 \tag{9.250}$$

which is guaranteed by

$$2\sigma_0 + \sigma_1 \leqslant 2 - d_s^+ - d_u^+ - 8\tau. \tag{9.251}$$

Then, as $\varepsilon_0^{\rho_1/3} (\log |\widetilde{I}|^{-1})^{3\tau^{-1}} |\widetilde{I}|^{\tau^2} \ll 1$ the required estimate holds for $\widehat{B}_3$.

9.12.3. Next consider $\widehat{B}_1$. We use $(\mathrm{SR1})_s$, $(\mathrm{SR1})_u$ to control the sizes of $\mathcal{C}_+(\widetilde{I})$, $\mathcal{C}_-(\widetilde{I})$, and (9.245) to estimate $(\rho_0 - d_s^*)^{-1}$. This gives

$$\begin{aligned} &\widehat{B}_1 \varepsilon_0^{A\tau} |\widetilde{I}|^{\tau(2-d_s^+-d_u^+-6\tau)} \\ &\quad \leqslant C \log^2 |\widetilde{I}|^{-1} (\rho_0 - d_s^*)^{-2} \left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{2\sigma} \varepsilon_0^{\tau(A-d_s^0-d_u^0)} |\widetilde{I}|^{\tau(2-d_s^+-d_u^+-6\tau)} \\ &\quad \leqslant C\tau^4 \log^2 \varepsilon_0^{-1} \log^2 |\widetilde{I}|^{-1} \left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{2\sigma+\tau(2-d_s^+-d_u^+-6\tau)} \varepsilon_0^{\tau(A-d_s^0-d_u^0+2-d_s^+-d_u^+-6\tau)}. \end{aligned} \tag{9.252}$$

On the other hand, in the range $x \leqslant \widetilde{x}_{cr}$, B_0 is smaller at $\widetilde{x}_{cr}$ where it is equal to

$$\begin{aligned} \left(\frac{\widetilde{x}_{cr}}{\varepsilon_0 |P_u|}\right)^{-\rho_0} \left(\frac{|I|}{\varepsilon_0}\right)^{2\sigma_0+\sigma_1} &= \left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{-\frac{\rho_0\sigma_0}{\rho_0-\rho_1}} \left(\frac{|I|}{\varepsilon_0}\right)^{2\sigma_0+\sigma_1} \\ &= \left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{-\frac{\rho_0\sigma_0}{\rho_0-\rho_1}+(1+\tau)(2\sigma_0+\sigma_1)} \varepsilon_0^{\tau(2\sigma_0+\sigma_1)}. \end{aligned} \tag{9.253}$$

We ask that

$$-\frac{\rho_0\sigma_0}{\rho_0-\rho_1} + (1+\tau)(2\sigma_0+\sigma_1) \leqslant 2\sigma + \tau(2 - d_s^+ - d_u^+ - 6\tau) - \tau^2, \tag{9.254}$$

and

$$A - d_s^0 - d_u^0 + 2 - d_s^+ - d_u^+ - 6\tau \geqslant 2\sigma_0 + \sigma_1 + 1. \tag{9.255}$$

Then, we will have, for all $x_{\min} \leqslant x \leqslant \widetilde{x}_{cr}$

$$\begin{aligned} \widehat{B}_1 \varepsilon_0^{A\tau} |\widetilde{I}|^{\tau(2-d_s^+-d_u^+-6\tau)} &\leqslant C\tau^4 \log^2 \varepsilon_0^{-1} \log^2 |\widetilde{I}|^{-1} \left(\frac{|\widetilde{I}|}{\varepsilon_0}\right)^{\tau^2} \varepsilon_0^{\tau} B_0 \\ &\ll B_0. \end{aligned} \tag{9.256}$$

9.12.4. In the case of $\widehat{B}_2$, we use $(SR1)_s$ to control the size of $\mathcal{C}_+(\widetilde{I})$ and $(SR2)_u$ to control $\sum_{\mathcal{C}_-(\widetilde{I})}(\max_{\widetilde{I}}|P_\omega|)^{\rho_s}$. This gives

$$\widehat{B}_2\varepsilon_0^{A\tau}|\widetilde{I}|^{\tau(2-d_s^+-d_u^+-6\tau)}$$
$$\leqslant C\Big(\frac{\log|\widetilde{I}|^{-1}}{\log\varepsilon_0^{-1}}\Big)^{2\tau^{-1}} x^{-\rho_s}\Big(\frac{|\widetilde{I}|}{\varepsilon_0}\Big)^{\sigma+\sigma_s}|P_u|^{\rho_s}\varepsilon_0^{\tau(A-d_s^0)}|\widetilde{I}|^{(1+\tau)\rho_s+\tau(2-d_s^+-d_u^+-6\tau)}$$
$$\leqslant C\Big(\frac{\log|\widetilde{I}|^{-1}}{\log\varepsilon_0^{-1}}\Big)^{2\tau^{-1}}\Big(\frac{x}{\varepsilon_0|P_u|}\Big)^{-\rho_s}\Big(\frac{|\widetilde{I}|}{\varepsilon_0}\Big)^{(1+\tau)\rho_s+\tau(2-d_s^+-d_u^+-6\tau)+\sigma+\sigma_s}\varepsilon_0^{A'\tau}, \tag{9.257}$$

with

$$A' := A+\rho_s-d_s^0+2-d_s^+-d_u^+-6\tau. \tag{9.258}$$

We want to show that the right-hand side of (9.257) is smaller than

$$B_0 = \Big(\frac{x}{\varepsilon_0|P_u|}\Big)^{-\rho_0}\Big(\frac{|I|}{\varepsilon_0}\Big)^{2\sigma_0+\sigma_1}. \tag{9.259}$$

As $\rho_0 > \rho_s$, it is sufficient to check this at $\widetilde{x}_{cr}$. After multiplying by $(\frac{\widetilde{x}_{cr}}{\varepsilon_0|P_u|})^{\rho_s}$, we have to show that

$$C\Big(\frac{\log|\widetilde{I}|^{-1}}{\log\varepsilon_0^{-1}}\Big)^{2\tau^{-1}}\Big(\frac{|\widetilde{I}|}{\varepsilon_0}\Big)^{(1+\tau)\rho_s+\tau(2-d_s^+-d_u^+-6\tau)+\sigma+\sigma_s}\varepsilon_0^{A'\tau} \tag{9.260}$$

is smaller than

$$\Big(\frac{|\widetilde{I}|}{\varepsilon_0}\Big)^{\frac{\sigma_0(\rho_s-\rho_0)}{\rho_0-\rho_1}+(1+\tau)(2\sigma_0+\sigma_1)}\varepsilon_0^{\tau(2\sigma_0+\sigma_1)}. \tag{9.261}$$

We will therefore ask that

$$\frac{\sigma_0(\rho_s-\rho_0)}{\rho_0-\rho_1}+(1+\tau)(2\sigma_0+\sigma_1)$$
$$\leqslant (1+\tau)\rho_s+\tau(2-d_s^+-d_u^+-6\tau)+\sigma+\sigma_s-\tau^2, \tag{9.262}$$

and

$$A+\rho_s-d_s^0+2-d_s^+-d_u^+-6\tau \geqslant 2\sigma_0+\sigma_1+1. \tag{9.263}$$

Then the required estimate is a consequence of

$$C\Big(\frac{\log|\widetilde{I}|^{-1}}{\log\varepsilon_0^{-1}}\Big)^{2\tau^{-1}}\Big(\frac{|\widetilde{I}|}{\varepsilon_0}\Big)^{\tau^2}\varepsilon_0^{\tau} \ll 1. \tag{9.264}$$

9.12.5. *Requirements on the exponents.* — Let us recall the various conditions imposed in this section on the various exponents. In Section 8, we introduced $d_s^* = d_s^0 + \varepsilon_0^{\frac{1}{5}d_s^0}$, $d_u^* = d_u^0 + \varepsilon_0^{\frac{1}{5}d_u^0}$, $d_s^+ = d_s^0 + C\eta\tau^{-1}$, $d_u^+ = d_u^0 + C\eta\tau^{-1}$.

– The exponent σ in (SR1) was defined in (9.49), Section 9.5, by

$$\sigma = \min(1 - d_u^+ - (1+\tau)d_s^* - \tau - c\eta\tau^{-1}, \quad 1 - d_s^+ - (1+\tau)d_u^* - \tau - c\eta\tau^{-1}). \tag{9.265}$$

– The exponents σ_s, σ_u in (SR2) were defined in (9.63), Section 9.6.1, by

$$\sigma_s := 1 - 3\tau - d_s^+, \qquad \sigma_u := 1 - 3\tau - d_u^+. \tag{9.266}$$

– The exponents ρ_s, ρ_u in (SR2) were asked to be close to d_s^0, d_u^0 respectively, to satisfy in (9.38), Section 9.4

$$\rho_u > d_u^0 + C\varepsilon_0, \qquad \rho_s > d_s^0 + C\varepsilon_0, \tag{9.267}$$

and later in (9.62) (Section 9.6.1) the stronger condition

$$\rho_u > d_u^* + \varepsilon_0^\tau, \qquad \rho_s > d_s^* + \varepsilon_0^\tau. \tag{9.268}$$

– The exponent ρ_0 in $(SR3)_s$ was required to be close to d_s^0. We asked also that $\rho_s > d_s^0 + C\varepsilon_0$ (cf. (9.38), Section 9.4), $\rho_0 > \rho_s + \varepsilon_0^\tau$ (cf. (9.62), Section 9.6.1), $\rho_0 \geqslant d_s^* + \varepsilon_0^{\tau^2/2}$ (cf. (9.98), Section 9.7.3), and finally (cf. (9.195), Section 9.10)

$$\rho_0 > d_s^* + C\tau^{-2}\frac{\log\log\varepsilon_0^{-1}}{\log\varepsilon_0^{-1}}. \tag{9.269}$$

– The exponent ρ_1 in $(SR3)_s$ was required to be close to $d_s^0(d_u^0 + 2d_s^0 - 1)\times (d_s^0 + d_u^0)^{-1}$.
– The exponents ρ_0', ρ_1' were defined by $\rho_0' = \frac{d_u^0}{d_s^0}\rho_0$, $\rho_1' = \frac{d_u^0}{d_s^0}\rho_1$.
– The exponents σ_0, σ_1 in (SR3) were required to be close to $1 - d_s^0$, $d_s^0 - d_u^0$ respectively. In (9.97), Section 9.7.3, we asked that

$$\sigma_0 + \sigma_1 \leqslant 1 - 3\tau - d_u^+, \tag{9.270}$$

and in (9.106), Section 9.7.4, that

$$\sigma_0 \leqslant 1 - 3\tau - d_s^+. \tag{9.271}$$

Finally, in (9.251), (9.254), (9.262) above, we required that

$$(\mathbf{9.272})\qquad \begin{aligned} &2\sigma_0+\sigma_1 \leqslant 2-d_s^+-d_u^+-8\tau,\\ &-\frac{\rho_0\sigma_0}{\rho_0-\rho_1}+(1+\tau)(2\sigma_0+\sigma_1)\leqslant 2\sigma+\tau(2-d_s^+-d_u^+-6\tau)-\tau^2,\\ &\frac{\sigma_0(\rho_s-\rho_0)}{\rho_0-\rho_1}+(1+\tau)(2\sigma_0+\sigma_1)\\ &\qquad \leqslant (1+\tau)\rho_s+\tau(2-d_s^+-d_u^+-6\tau)+\sigma+\sigma_s-\tau^2. \end{aligned}$$

– The exponent A in (SR3) was required above to satisfy

$$(\mathbf{9.273})\qquad A-d_s^0-d_u^0+2-d_s^+-d_u^+-6\tau \geqslant 2\sigma_0+\sigma_1+1,$$

$$(\mathbf{9.274})\qquad A+\rho_s-d_s^0+2-d_s^+-d_u^+-6\tau \geqslant 2\sigma_0+\sigma_1+1.$$

We have proved that, if all these conditions are satisfied, all candidates but a proportion not larger than $C|\widetilde{I}|^{\tau^2}$ satisfy $(SR3)_s$. Before checking that these requirements on the exponents are compatible, we need to review (briefly) the induction step for $(SR3)_u$ because new requirements will appear from it.

9.13. *The induction step for* $(SR3)_u$. — The proof that most candidates I in a strongly regular parent $\widetilde{I}$ satisfy $(SR3)_u$ follows the same plan than for $(SR3)_s$. However, condition (SR3) is not symmetric, hence we must check that the various steps work in the same way. We review briefly these steps below.

– *Very small values of x* (Section 9.7.1): we use now

$$(\mathbf{9.275})\qquad \#Bi_-(I, I_\alpha, I_\omega; x) \leqslant C x^{-d_u^*},$$

for $x \leqslant x'_{\min} := |\widetilde{I}|^{C(\rho_0'-d_u^*)^{-1}}$.

– *Old and new elements* (Section 9.7.2) are defined as before.

– Observe that σ_0, σ_1 play the same role in $(SR3)_u$ than in $(SR3)_s$. Therefore, the results in Sections 9.7.3, 9.7.4, 9.7.5 work exactly in the same way, replacing $\widetilde{x}_{cr}$ in Proposition 42 by $\widetilde{x}'_{cr} := \varepsilon_0|Q_s|(\frac{|\widetilde{I}|}{\varepsilon_0})^{\frac{\sigma_0}{\rho_0'-\rho_1'}}$. We require as in (9.98) that $\rho_0' \geqslant d_u^* + \varepsilon_0^{\tau^2/2}$, but this is actually a consequence of the definition of ρ_0' and of (9.269) above.

– *Size of* Bi^{new} (Section 9.8). One proceeds as in Sections 9.8.1, 9.8.2, separating then the sum corresponding to (9.139) into two parts S'_l and S'_s. In the first sum (Section 9.8.3), the exponent ρ_1 must be replaced by $\rho_1' = \rho_1\frac{d_u^0}{d_s^0} \leqslant \rho_1$. In particular the exponent $\sigma_0+\sigma_1-\rho_1'$ in (9.146) is still positive. The same considerations apply for S'_s. In the case $d_s^0 = d_u^0 = 1/2$, the exponent σ'_2 corresponding to σ_2 now depends on ω_s.

– Except for the obvious modifications in the definitions of the $b^+(P_\alpha, \ell)$, $b^-(Q_\omega, \ell)$, Section 9.9 is unchanged. In particular, the set L is defined in the same way; to estimate its cardinality, we now assume

$$\rho'_0 > d_u^* + C\tau^{-2}\frac{\log\log\varepsilon_0^{-1}}{\log\varepsilon_0^{-1}}, \tag{9.276}$$

but, in view of the definition of $\rho'_0 = \rho_0 d_u^0/d_s^0$, this is the same thing than (9.269), up to the value of the constant C.

– The quantities $\widehat{B}'_i$, $i = 1, 2, 3$ that appear instead of the $\widehat{B}_i$ of Proposition 46 are

$$\widehat{B}'_1 = (\#\mathcal{C}_+(\widetilde{I}))(\#\mathcal{C}_-(\widetilde{I}))(\log|\widetilde{I}|^{-1})^2(\rho'_0 - d_u^*)^{-2},$$

$$\widehat{B}'_2 = \left(\frac{\log|\widetilde{I}|^{-1}}{\log\varepsilon_0^{-1}}\right)^{2\tau^{-1}}(C|\widetilde{I}|^{1+\tau}x^{-1})^{\rho_u}(\#\mathcal{C}_-(\widetilde{I}))\sum_{\mathcal{C}_+(\widetilde{I})}\left(\max_{\widetilde{I}}|Q_\alpha|\right)^{\rho_u},$$

$$\widehat{B}'_3 = \varepsilon_0^{\rho'_1/3}|\widetilde{I}|^{-\varepsilon_0^{\rho'_1/3}}(\log|\widetilde{I}|^{-1})^{3\tau^{-1}}\widetilde{B}'.$$

– Dealing with $\widehat{B}'_3$ requires (9.251) as for $\widehat{B}_3$.

– Dealing with $\widehat{B}'_1$ requires (9.254), (9.255) as for $\widehat{B}_1$. Indeed, observe that in (9.254), we have

$$\frac{\rho_0}{\rho_0 - \rho_1} = \frac{\rho'_0}{\rho'_0 - \rho'_1} \tag{9.277}$$

in view of the proportionality of the exponents by the factor d_u^0/d_s^0.

– To control $\widehat{B}'_2$, we now require, corresponding to (9.262), (9.263)

$$\frac{\sigma_0(\rho_u - \rho'_0)}{\rho'_0 - \rho'_1} + (1+\tau)(2\sigma_0 + \sigma_1)$$
$$\leqslant (1+\tau)\rho_u + \tau(2 - d_s^+ - d_u^+ - 6\tau) + \sigma + \sigma_u - \tau^2, \tag{9.278}$$
$$A + \rho_u - d_u^0 + 2 - d_s^+ - d_u^+ - 6\tau \geqslant 2\sigma_0 + \sigma_1 + 1. \tag{9.279}$$

9.14. *Conclusion.*

Theorem **3.** — *Assume that the exponents in* (SR1), (SR2), (SR3) *satisfy the requirements of Section 9.12.5 and also (9.278), (9.279). Then, all candidates* I *in a strongly regular interval* $\widetilde{I}$ *are strongly regular, except for a proportion not larger than* $C|\widetilde{I}|^{\tau^2}$. *Moreover, it is possible to choose the exponents in order to satisfy these assumptions.*

Proof. — The first statement has been proved in Sections 9.5 through 9.13!

We choose the undefined exponents (σ, σ_s, σ_u are already defined) in the following order. First, take $\rho_u = d_u^* + 2\varepsilon_0^\tau$, $\rho_s = \frac{d_s^0}{d_u^0}\rho_u$. Then (9.268) is satisfied, and thus also (9.267). Next, choose

$$\rho_1 = d_s^0(d_u^0 + 2d_s^0 - 1)(d_s^0 + d_u^0)^{-1},$$
$$\rho_1' = d_u^0(d_u^0 + 2d_s^0 - 1)(d_s^0 + d_u^0)^{-1},$$
$$\rho_0 = d_s^0\left(1 + \tau^{-3}\frac{\log\log\varepsilon_0^{-1}}{\log\varepsilon_0^{-1}}\right),$$
$$\rho_0' = d_u^0\left(1 + \tau^{-3}\frac{\log\log\varepsilon_0^{-1}}{\log\varepsilon_0^{-1}}\right).$$

Then all the requirements on the ρ exponents are satisfied.

The exponents σ_0, σ_1 must satisfy (9.270), (9.271), the three relations in (9.272), and (9.278). Each of this six relations defines in the σ_0, σ_1-plane an affine half-plane whose boundary passes very close to the point $(1 - d_s^0, 1 - d_u^0)$ which has to be close to (σ_0, σ_1). Observe first that the coefficient of σ_1 in each of the six relations is positive. This implies that the intersection of the 6 *linear* half-planes parallel to these 6 affine half-planes contains an open cone at the origin. But then it is possible to find (σ_0, σ_1) close to $(1 - d_s^0, 1 - d_u^0)$ in the intersection of the affine half-planes.

Finally, the constraints on A are clearly satisfied if A is large enough. Actually, A = 3 works! □

Definition **13.** — *A strongly regular parameter (in* I_0*) is one which is the intersection of a decreasing sequence of strongly regular parameter intervals.*

Corollary **15.** — *Except for a set of relative measure* $\leqslant C\varepsilon_0^{\tau^2}$*, parameters in* I_0 *are strongly regular.*

Proof. — By Proposition 31, the starting interval is strongly regular. From the theorem, in a strongly regular parameter interval of level k, all points are contained in a strongly regular interval of the next level except for a set of relative Lebesgue measure $\leqslant \varepsilon_k^{\tau^2}$. As we have

$$\sum_{k\geqslant 0}\varepsilon_k^{\tau^2} \leqslant C\varepsilon_0^{\tau^2}, \tag{9.280}$$

the statement of the Corollary follows. □

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 $(I_m)_{m\geqslant 0}$ of strongly regular parameter intervals.

The sequence $\mathcal{R}(I_m)$ is increasing and we set

$$\mathcal{R} = \bigcup_{m\geqslant 0} \mathcal{R}(I_m). \tag{10.1}$$

Definition **14.** — *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.

Lemma **14.** — *Let* $(P, Q, n) \in \mathcal{R}$ *be an element which can be written as a simple composition* $(P, Q, n) = (P_1, Q_1, n_1) * (P_2, Q_2, n_2)$. *Let* $(\widetilde{P}_2, \widetilde{Q}_2, \widetilde{n}_2)$ *be the element such that* $\widetilde{P}_2$ *is the parent of* P_2, *and let* $(\widetilde{P}, \widetilde{Q}, \widetilde{n}) := (P_1, Q_1, n_1) * (\widetilde{P}_2, \widetilde{Q}_2, \widetilde{n}_2)$. *Then* $\widetilde{P}$ *is the parent of* P.

Proof. — If P_2 is a simple child, we have $n_2 = \widetilde{n}_2 + 1$, $n = \widetilde{n} + 1$ and P is a simple child of $\widetilde{P}$.

We now assume that P_2 is a non-simple child of $\widetilde{P}_2$. Let $(P', Q', n') \in \mathcal{R}$ the element such that P' is the child of $\widetilde{P}$ containing P. We will show that $P' = P$. We have $g_t^{\widetilde{n}}(P) \subset g_t^{\widetilde{n}_2}(P_2) \subset L_u$, hence P' also is a non-simple child.

Applying twice Proposition 7 (Section 6.2), we can write, in some $\mathcal{R}(I_m)$

$$(P_2, Q_2, n_2) \in (\widetilde{P}_2, \widetilde{Q}_2, \widetilde{n}_2) \,\Box\, (P_3, Q_3, n_3),$$
$$(P', Q', n') \in (\widetilde{P}, \widetilde{Q}, \widetilde{n}) \,\Box\, (P'_3, Q'_3, n'_3).$$

We have $\widetilde{Q} \subset \widetilde{Q}_2$ and also $P_3 \subset P'_3$ because $P \subset P'$. As $\widetilde{Q}_2 \pitchfork_{I_m} P_3$ and $\widetilde{Q} \pitchfork_{I_m} P'_3$ hold, concavity (Proposition 9, Section 6.3) imply that $\widetilde{Q}_2 \pitchfork_{I_m} P'_3$ also holds. As P_2 is a child of $\widetilde{P}_2$, we must have $P_3 = P'_3$ and $P' = P$. □

Proposition **47.** — *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

$$(\mathrm{P}, \mathrm{Q}, n) = (\mathrm{P}_1, \mathrm{Q}_1, n_1) * \cdots * (\mathrm{P}_r, \mathrm{Q}_r, n_r) \tag{10.2}$$
$$= (\mathrm{P}'_1, \mathrm{Q}'_1, n'_1) * \cdots * (\mathrm{P}'_s, \mathrm{Q}'_s, n'_s).$$

It is sufficient to show that $(\mathrm{P}_1, \mathrm{Q}_1, n_1) = (\mathrm{P}'_1, \mathrm{Q}'_1, n'_1)$. This is true if $n_1 = n'_1$. Assume for instance that $n_1 < n'_1$. Then we have $\mathrm{P} \subset \mathrm{P}'_1 \subset \mathrm{P}_1$ with $\mathrm{P}'_1 \neq \mathrm{P}_1$. But a repeated application of the lemma above shows that all rectangles in-between P and P_1 can be written as simple compositions of $(\mathrm{P}_1, \mathrm{Q}_1, n_1)$ with some other element. In particular, this is the case for $(\mathrm{P}'_1, \mathrm{Q}'_1, n'_1)$, which cannot be prime. □

Remark **14.** — In the prime decomposition

$$(\mathrm{P}, \mathrm{Q}, n) = (\mathrm{P}_1, \mathrm{Q}_1, n_1) * \cdots * (\mathrm{P}_r, \mathrm{Q}_r, n_r), \tag{10.3}$$

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 $(\mathrm{P}, \mathrm{Q}, n)$ be an element of $\mathcal{R}^*$ and let

$$(\mathrm{P}, \mathrm{Q}, n) = (\mathrm{P}_1, \mathrm{Q}_1, n_1) * \cdots * (\mathrm{P}_r, \mathrm{Q}_r, n_r), \tag{10.4}$$

be its prime decomposition. We define

$$\mathrm{T}^+((\mathrm{P}, \mathrm{Q}, n)) = (\mathrm{P}_2, \mathrm{Q}_2, n_2) * \cdots * (\mathrm{P}_r, \mathrm{Q}_r, n_r), \tag{10.5}$$
$$\mathrm{T}^-((\mathrm{P}, \mathrm{Q}, n)) = (\mathrm{P}_1, \mathrm{Q}_1, n_1) * \cdots * (\mathrm{P}_{r-1}, \mathrm{Q}_{r-1}, n_{r-1}),$$

if $r > 1$. When $(\mathrm{P}, \mathrm{Q}, n)$ is prime, with $\mathrm{P} \subset \mathrm{R}_a$ and $\mathrm{Q} \subset \mathrm{R}_{a'}$, we set

$$\mathrm{T}^+((\mathrm{P}, \mathrm{Q}, n)) = (\mathrm{R}_{a'}, \mathrm{R}_{a'}, 0), \tag{10.6}$$
$$\mathrm{T}^-((\mathrm{P}, \mathrm{Q}, n)) = (\mathrm{R}_a, \mathrm{R}_a, 0).$$

For $\mathrm{S} = (\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}$, we write $\mathrm{S} * \mathcal{R}$, resp. $\mathcal{R} * \mathrm{S}$, for the set of elements which can be written as $(\mathrm{P}, \mathrm{Q}, n) * (\mathrm{P}', \mathrm{Q}', n')$, resp. $(\mathrm{P}', \mathrm{Q}', n') * (\mathrm{P}, \mathrm{Q}, n)$, for some $(\mathrm{P}', \mathrm{Q}', n') \in \mathcal{R}$. We have partitions

$$\mathcal{R}^* = \bigsqcup_{\mathcal{P}} \mathrm{S} * \mathcal{R} = \bigsqcup_{\mathcal{P}} \mathcal{R} * \mathrm{S}. \tag{10.7}$$

10.2. *Number of factors in a prime decomposition.* — We write $r(\mathrm{S})$ for the number of factors in the prime decomposition of an element S of $\mathcal{R}$ (setting $r(\mathrm{S}) = 0$ if S has length 0). Let $(\mathrm{P}, \mathrm{Q}, n)$, $(\mathrm{P}', \mathrm{Q}', n')$ 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

$$r(\mathrm{P}', \mathrm{Q}', n') = r(\mathrm{P}, \mathrm{Q}, n) + 1. \tag{10.8}$$

On the other hand, assume that P′ is a non-simple child of P. Let m be such that $(P', Q', n') \in \mathcal{R}(I_m)$. By Proposition 7 (Section 6.2), we can write

$$(\mathbf{10.9}) \qquad (P', Q', n') \in (P, Q, n) \,\square\, (\widehat{P}, \widehat{Q}, \widehat{n}),$$

for some $(\widehat{P}, \widehat{Q}, \widehat{n}) \in \mathcal{R}(I_m)$ such that $Q \pitchfork_{I_m} \widehat{P}$ holds. Let

$$(\mathbf{10.10}) \qquad (P, Q, n) = (P_1, Q_1, n_1) * \cdots * (P_r, Q_r, n_r)$$

be the prime decomposition of (P, Q, n).

Define, for $1 \leqslant i \leqslant r$

$$\begin{aligned} (P^i, Q^i, n^i) &= (P_i, Q_i, n_i) * \cdots * (P_r, Q_r, n_r) \\ (\mathbf{10.11}) \qquad &= (T^+)^{i-1}(P, Q, n). \end{aligned}$$

We have an increasing sequence

$$(\mathbf{10.12}) \qquad Q = Q^1 \subset Q^2 \subset \cdots \subset Q^r = Q_r.$$

Let r' be the largest integer in $\{1, \ldots, r\}$ such that $Q^{r'}$ and $\widehat{P}$ are I_l-transverse for some $l \geqslant m$ (and then for all large enough l). Define $(\widetilde{P}, \widetilde{Q}, \widetilde{n}) \in \mathcal{R}$ by the condition $Q' \subset \widetilde{Q}$ and

$$(\mathbf{10.13}) \qquad (\widetilde{P}, \widetilde{Q}, \widetilde{n}) \in (P^{r'}, Q^{r'}, n^{r'}) \,\square\, (\widehat{P}, \widehat{Q}, \widehat{n}).$$

Proposition **48.** — *The simple composition*

$$(\mathbf{10.14}) \qquad (P', Q', n') = (P_1, Q_1, n_1) * \cdots * (P_{r'-1}, Q_{r'-1}, n_{r'-1}) * (\widetilde{P}, \widetilde{Q}, \widetilde{n}).$$

is the decomposition of (P', Q', n') *in prime factors. In particular, we have*

$$r(P', Q', n') = r' \leqslant r(P, Q, n).$$

Proof. — It is sufficient to show that $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$ is prime. Assume by contradiction that we can write

$$(\mathbf{10.15}) \qquad (\widetilde{P}, \widetilde{Q}, \widetilde{n}) = (\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1) * (\widetilde{P}_2, \widetilde{Q}_2, \widetilde{n}_2)$$

with $\widetilde{n}_1, \widetilde{n}_2 > 0$. Define

$$(\mathbf{10.16}) \qquad (P'_1, Q'_1, n'_1) = (P_1, Q_1, n_1) * \cdots * (P_{r'-1}, Q_{r'-1}, n_{r'-1}) * (\widetilde{P}_1, \widetilde{Q}_1, \widetilde{n}_1).$$

Claim. — *There exists* $r' \leqslant j \leqslant r$ *such that*

$$(\mathbf{10.17}) \qquad (P'_1, Q'_1, n'_1) = (P_1, Q_1, n_1) * \cdots * (P_j, Q_j, n_j).$$

Proof of claim. — We have $P' \subset P'_1$, $P' \neq P'_1$, hence $P \subset P'_1$, $n'_1 \leqslant n$. As $\widetilde{n}_1 > 0$, the smallest integer j such that

$$(\mathbf{10.18}) \qquad n'_1 \leqslant n_1 + \cdots + n_j,$$

satisfies $r' \leqslant j \leqslant r$. From

$$(\mathbf{10.19}) \qquad (P', Q', n') = (P'_1, Q'_1, n'_1) * (\widetilde{P}_2, \widetilde{Q}_2, \widetilde{n}_2)$$

and Lemma 14, we can write $(P_1, Q_1, n_1) * \cdots * (P_j, Q_j, n_j)$ as a simple composition

$$(\mathbf{10.20}) \qquad (P_1, Q_1, n_1) * \cdots * (P_j, Q_j, n_j) = (P'_1, Q'_1, n'_1) * (\overline{P}, \overline{Q}, \overline{n})$$

for some $(\overline{P}, \overline{Q}, \overline{n}) \in \mathcal{R}$. Still by Lemma 14, we can write

$$(\mathbf{10.21}) \qquad (P_j, Q_j, n_j) = (\overline{P}', \overline{Q}', \overline{n}') * (\overline{P}, \overline{Q}, \overline{n})$$

for some $(\overline{P}', \overline{Q}', \overline{n}') \in \mathcal{R}$. We have $\overline{n}' > 0$ by definition of j. As (P_j, Q_j, n_j) is prime, we have $\overline{n} = 0$, which proves the claim

We will now prove that, for l large enough, $Q^{j+1} \pitchfork_{I_l} \widehat{P}$ holds, a contradiction with the definition of r'. Indeed, from (10.19) we have $\widetilde{P}_2 \subset P^{j+1}$, $\widetilde{P}_2 \neq P^{j+1}$. Let $(P^*, Q^*, n^*) \in \mathcal{R}$ the element such that P^* is the child of P^{j+1} containing $\widetilde{P}_2$. By Proposition 7 (Section 6.2), we can write (with l large enough) $(P^*, Q^*, n^*) \in (P^{j+1}, Q^{j+1}, n^{j+1}) \,\square\, (\widehat{P}^*, \widehat{Q}^*, \widehat{n}^*)$, for some $(\widehat{P}^*, \widehat{Q}^*, \widehat{n}^*)$ with $\widehat{P}^* \supset \widehat{P}$. Then $Q^{j+1} \pitchfork_{I_l} \widehat{P}^*$ holds and $Q^{j+1} \pitchfork_{I_l} \widehat{P}$ also holds.

Thus, we obtain a contradiction and the proof of the proposition is complete. □

10.3. *A weighted estimate on the number of children.* — We present in this subsection a variation over the estimates in Section 8.3, which will be important in the definition of a transfer operator.

As mentioned already in Section 9.4, there exist from classical results of Bowen, Ruelle and Sinai an invariant probability measure m^+ on K such that the measure of an I_0-defined rectangle P is of the order $|P|^{d_s^t}$, where d_s^t is the transverse Hausdorff dimension of the stable foliation $W^s(K)$. Recall that d_s^t is a smooth, hence Lipschitz, function of t.

We fix a constant $\kappa \in (0,1)$ close to 1, but independent of ε_0. Let also $d_s^- < d_s^t$, close to d_s^t.

For $S = (P, Q, n)$, we set

$$(\mathbf{10.22}) \qquad \|P\| = |P|^{d_s^-} \kappa^{r(S)}$$

(we will also write $r(P)$ instead of $r(S)$).

Proposition **49.** — *Assume that* $\log \kappa^{-1} < C^{-1}$ *and* $d_s^t - d_s^- < C^{-1} \log \kappa^{-1}$ *with* C *large enough. For any* $m \geqslant 1$*, any* $(P, Q, n) \in \mathcal{R}$*, we have*

$$\sum_{P'} \|P'\| \leqslant C\kappa^{\frac{m}{2}} \|P\|$$

where the sum in the left-hand side is over elements (P', Q', n') *such that* P' *is a descendant 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 **15.** — *Let* $\varepsilon_1 > 0$*. If* ε_0 *is small enough, we have*

$$\sum_{P'} \|P'\| \leqslant \varepsilon_1 \|P\|$$

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 $m_0 \geqslant 1$ be an integer to be determined later. Consider all chains

$$P = P^0 \supset P^1 \supset \cdots \supset P^{m_0} = P' \tag{10.23}$$

where P is given and P^{i+1} is a child of P^i.

First consider the case where P^{i+1} is, for each i, a simple child of P^i. One has then $r(P') = m_0 + r(P)$, and one can write $(P', Q', n') = (P, Q, n) * (\widehat{P}, \widehat{Q}, \widehat{n})$ with $(\widehat{P}, \widehat{Q}, \widehat{n}) \in \mathcal{R}(I_0)$. We have

$$\begin{aligned}
\|P'\| &\leqslant C\kappa^{m_0} \|P\| |\widehat{P}|^{d_s^-} \\
&\leqslant C\kappa^{m_0} \|P\| m^+(\widehat{P}) |\widehat{P}|^{d_s^- - d_s^t} \\
&\leqslant C\kappa^{m_0} \|P\| m^+(\widehat{P}) \exp(Cm_0(d_s^t - d_s^0)) \\
&\leqslant C\kappa^{\frac{2}{3}m_0} \|P\| m^+(\widehat{P}),
\end{aligned}$$

if $d_s^t - d_s^- < C^{-1} \log \kappa^{-1}$ with C large enough. Under this condition the part of the sum in Proposition 49 corresponding to simple descendants satisfies (as $\sum m^+(\widehat{P}) \leqslant 1$)

$$\sum \|P'\| \leqslant C\kappa^{\frac{2}{3}m_0} \|P\|. \tag{10.24}$$

We choose m_0 such that in (10.24) we have

$$C\kappa^{\frac{2}{3}m_0} \leqslant \frac{1}{2} \kappa^{\frac{m_0}{2}}. \tag{10.25}$$

On the other hand, from the lemma above, it follows that for every $\widetilde{P}$, we have

$$(\mathbf{10.26})\qquad \sum \|\widetilde{P}'\| \leqslant C\|\widetilde{P}\|$$

where the sum is over all children of $\widetilde{P}$. Using the lemma again, when we sum over chains such that P_{i+1} is a non-simple child of P_i for at least one i, we obtain

$$(\mathbf{10.27})\qquad \sum \|\widetilde{P}'\| \leqslant m_0 C^{m_0-1}\varepsilon_1 \|P\|.$$

Taking ε_1 small enough, we obtain

$$(\mathbf{10.28})\qquad \sum \|P'\| \leqslant \kappa^{\frac{m_0}{2}} \|P\|$$

where the sum is now over all chains. The proposition follows immediately from (10.28) and (10.26). □

Lemma **16.** — *Let* $(P_0, Q_0, n_0), (P'_0, Q'_0, n'_0), (P_1, Q_1, n_1) \in \mathcal{R}$ *with* $Q_0 \subset Q'_0$. *If* $Q_0 \pitchfork_{I_m} P_1$ *holds for m large enough and*

$$8|Q'_0|^{1-\eta} < \delta(Q_0, P_1),$$

then $Q'_0 \pitchfork_{I_m} P_1$ *holds for m large enough.*

Proof. — If $3|P_1|^{1-\eta} < \delta(Q_0, P_1)$, $Q'_0 \overline{\pitchfork}_{I_m} P_1$ holds for m large enough by direct verification of (T1), (T2), (T3) of Section 5.4. If on the other hand, $3|P_1|^{1-\eta} \geqslant \delta(Q_0, P_1)$, we can apply Proposition 21 in Section 8.1 to conclude that $Q'_0 \pitchfork_{I_m} P_1$ holds for m large enough. □

Proof of Lemma 15. — Let $(P, Q, n) \in \mathcal{R}$. Any non-simple child P' of P is obtained as

$$(\mathbf{10.29})\qquad (P', Q', n') \in (P, Q, n) \boxdot (P_1, Q_1, n_1)$$

and we denote by $\widetilde{P}_1$, the parent of P_1. One has

$$(\mathbf{10.30})\qquad |P'| \leqslant C|P||P_1|\delta(Q, P_1)^{-\frac{1}{2}}.$$

Therefore, we will have

$$(\mathbf{10.31})\qquad \|P\|^{-1} \sum \|P'\| \leqslant C \sum |P_1|^{d_s^-} \kappa^{r(P')-r(P)} \delta(Q, P_1)^{-\frac{1}{2}d_s^-}.$$

By Proposition 48, there is an increasing sequence

$$(\mathbf{10.32})\qquad Q = Q^1 \subset Q^2 \subset \cdots \subset Q^{r(P)} = Q_{r(P)}$$

such that $r(P')$ is the largest integer r' for which $Q^{r'}$ and P_1 are I_m-transverse for large enough m. We claim that

$$r(P) - r(P') \leqslant C \log(\delta(Q, P_1))^{-1}. \tag{10.33}$$

Indeed, let $1 \leqslant r \leqslant r(P)$ such that

$$|Q^r| \leqslant \delta(Q, P_1)^2. \tag{10.34}$$

By Lemma 16, $Q^r \pitchfork_{I_m} P_1$ holds for m large enough, hence $r \leqslant r(P)$. On the other hand, there exists $\kappa^* \in (0, 1)$ such that

$$|Q^r| \leqslant \kappa^* |Q^{r+1}| \tag{10.35}$$

for $r < r(P)$. The claim follows.

Therefore, if κ is close enough to 1, we have

$$\kappa^{r(P')-r(P)} \leqslant (\delta(Q, P_1))^{-\frac{1}{6}d_s^-} \tag{10.36}$$

and the right-hand side of (10.31) is bounded by

$$C \sum |P_1|^{d_s^-} (\delta(Q, P_1))^{-\frac{2}{3}d_s^-}. \tag{10.37}$$

Using (R7) (Section 5.4), this is smaller than

$$C \sum |P_1|^{\frac{1}{3}d_s^-}. \tag{10.38}$$

To estimate this sum, we first fix the parent $\widetilde{P}_1$ and sum over children P_1; it follows from Proposition 26 (Section 8.2) that the corresponding sum is bounded by $C|\widetilde{P}_1|^{\frac{1}{3}d_s^-}$. Then (10.38) is not greater than

$$C \sum |\widetilde{P}_1|^{\frac{1}{3}d_s^-}. \tag{10.39}$$

For each integer m, let us count now how many $\widetilde{P}_1$ may satisfy

$$2^{-m} \geqslant |\widetilde{P}_1| \geqslant 2^{-m-1}. \tag{10.40}$$

As P' is a child of P, Q is transverse to P_1 but not to $\widetilde{P}_1$. From Lemma 16, we have therefore

$$\delta(Q, \widetilde{P}_1) \leqslant 8|\widetilde{P}_1|^{1-\eta} \leqslant 8\, 2^{-m(1-\eta)} \tag{10.41}$$

which shows that there are no more than $C2^{m\eta}$ such $\widetilde{P}_1$'s. This implies that the sum (10.39) is at most of order $C\varepsilon_0^{\frac{1}{3}d_s^-}$, which yields the statement of the lemma. □

Remark **15.** — In Lemma 15, the value d_s^- 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.

Corollary **16.** — *Let* $\varepsilon_1 > 0$. *If* ε_0 *is small enough, we have*

$$\sum_{\mathcal{P}} n|P|^{d_s^-} < \varepsilon_1$$

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 12 in Section 6.6.2

$$n \leqslant \left(\log(C|P|^{-1})\right)^{\frac{\log 2}{\log 3/2}}. \tag{10.42}$$

Choose $\widehat{d_s^-} < d_s^-$ but so close to d_s^- that the hypothesis $d_s^t - \widehat{d_s^-} < C^{-1}\log\kappa^{-1}$ of Proposition 49 is still satisfied. If ε_0 is small enough, we have, for any non trivial prime P, as $|P| < \varepsilon_0$

$$n|P|^{d_s^-} \leqslant |P|^{\widehat{d_s^-}}. \tag{10.43}$$

Observe also that the thinnest $(\widetilde{P}, \widetilde{Q}, \widetilde{n}) \in \mathcal{R}(I_0)$ with $P \subset \widetilde{P}$ satisfies $\widetilde{Q} \subset Q_u$ hence $|\widetilde{P}| \leqslant \varepsilon_0^{\alpha}$ for some fixed positive α.

We apply the proposition using $\widehat{d_s^-}$ instead of d_s^-; we obtain, as $r(P) = 1$ for primes

$$\begin{aligned}\sum n|P|^{d_s^-} &\leqslant \sum |P|^{\widehat{d_s^-}}\\ &= \kappa^{-1}\sum \|P\|\\ &\leqslant C\kappa^{-1}(1-\kappa^{\frac{1}{2}})^{-1}\sum \|\widetilde{P}\|,\end{aligned} \tag{10.44}$$

where, in the last sum, $(\widetilde{P}, \widetilde{Q}, \widetilde{n})$ runs through the elements of $\mathcal{R}(I_0)$ with $|\widetilde{P}|$ of the order of ε_0^{α}. We have

$$\begin{aligned}\sum \|\widetilde{P}\| &\leqslant C\sum \varepsilon_0^{\alpha\widehat{d_s^-}}\kappa^{C^{-1}\log\varepsilon_0^{-1}}\\ &\leqslant C\varepsilon_0^{\alpha(\widehat{d_s^-}-d_s^t)+C^{-1}\log\kappa^{-1}},\end{aligned} \tag{10.45}$$

and the exponent is positive from the hypothesis of Proposition 49. Putting this into (10.44) yields Corollary 16. □

Remark **16.** — 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 is convergent for $d > d_s^0 - C^{-1}$, C large enough.

10.4. *Stable curves.* — Let $(P_k, Q_k, n_k)_{k\geqslant 0}$ be a sequence of elements of $\mathcal{R}$ such that P_{k+1} is strictly contained in P_k for $k \geqslant 0$. 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 $\{x_a = \varphi_k^\pm(y_a)\}$.

Proposition **50.** — *The intersection* $\bigcap_{k\geqslant 0} P_k$ *is the graph* $\{x_a = \varphi_\infty(y_a)\}$ *of a* $C^{1+\mathrm{Lip}}$ *function, whose* $C^{1+\mathrm{Lip}}$ *norm of* φ_∞ *is bounded independently of the sequence* $(P_k)_{k\geqslant 0}$. *Moreover, we have, for all* $k \geqslant 0, y_a \in I_a^u$

$$|\varphi_k^\pm(y_a) - \varphi_\infty(y_a)| \leqslant C|P_k|,$$
$$|D\varphi_k^\pm(y_a) - D\varphi_\infty(y_a)| \leqslant C|P_k|.$$

Proof. — Let (A_k, B_k) be the implicit representation associated to (P_k, Q_k, n_k). We have

$$\varphi_k^\pm(y_a) = A_k(y_a, x_b^\pm), \tag{10.46}$$

where $x_b^\pm$ are the endpoints of I_b^s and $Q_k \subset R_b$. The partial derivatives $A_{k,y}$, $A_{k,yy}$ are bounded by u_0^{-1}, $2D_0$ respectively and we have

$$|\varphi_k^+(y_a) - \varphi_k^-(y_a)| \leqslant C|P_k|. \tag{10.47}$$

This implies the statement of the proposition, except for the last inequality.

To compare the derivatives of $\varphi_k^\pm$ and $\varphi_{k+1}^\pm$, we use (A.66) in Appendix A if P_{k+1} is a simple child of P_k, (A.86) if it is a non-simple child. We obtain that, for every y_a, x, there exists x^* such that

$$|A_{k+1,y}(y_a, x) - A_{k,y}(y_a, x^*)| \leqslant C|P_k||Q_k|, \tag{10.48}$$

in the first case and

$$|A_{k+1,y}(y_a, x) - A_{k,y}(y_a, x^*)| \leqslant C|P_k||Q_k|\delta^{-1/2}, \tag{10.49}$$

in the second case. Here, we have $\delta \ll |Q_k|$ from (R7) in Section 5.4. Using that

$$|A_{k,y}(y_a, x^*) - A_{k,y}(y_a, x_b^\pm)| \leqslant C\|A_{k,xy}\|_\infty \leqslant C|P_k|, \tag{10.50}$$

we conclude that in both cases we have

$$|D\varphi_k^{\pm}(y_a) - D\varphi_{k+1}^{\pm}(y_a)| \leqslant C|P_k|. \tag{10.51}$$

As there exists $\kappa^* \in (0, 1)$ such that $|P_{k+1}| \leqslant \kappa^*|P_k|$ for all $k \geqslant 0$, we obtain, for all $l > k$

$$|D\varphi_k^{\pm}(y_a) - D\varphi_l^{\pm}(y_a)| \leqslant C|P_k|, \tag{10.52}$$

and the required inequality follows, letting l go to $+\infty$. □

Definition **15.**

1. *A stable curve is the intersection* $\omega = \bigcap_{k\geqslant 0} P_k$ *of a decreasing sequence of vertical-like rectangles as above. An unstable curve is the intersection* $\omega' = \bigcap_{k\geqslant 0} Q'_k$ *of a decreasing sequence of horizontal-like strips.*
2. *The set of stable curves, resp. unstable curves, is denoted by* $\mathcal{R}_+^\infty$*, resp.* $\mathcal{R}_-^\infty$*. The union of stable curves, resp. unstable curves, is denoted by* $\widetilde{\mathcal{R}}_+^\infty$*, resp.* $\widetilde{\mathcal{R}}_-^\infty$*.*
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*

$$\pi : \widetilde{\mathcal{R}}_+^\infty \mapsto \mathcal{R}_+^\infty.$$

We will now define dynamics on a part of the sets $\mathcal{R}_+^\infty$, $\widetilde{\mathcal{R}}_+^\infty$.

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 $\mathcal{R}_+^\infty$. 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

$$\mathcal{R}_+^\infty = \mathcal{N}_+ \bigsqcup \mathcal{D}_+, \tag{10.53}$$

$$\mathcal{D}_+ = \bigsqcup_{\mathcal{P}} \mathcal{R}_+^\infty(P). \tag{10.54}$$

We denote by $\widetilde{\mathcal{N}}_+$, $\widetilde{\mathcal{D}}_+$, $\widetilde{\mathcal{R}}_+^\infty(P)$ 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 47)

$$(P_k, Q_k, n_k) = (P, Q, n) * (P'_k, Q'_k, n'_k) \tag{10.55}$$

for some $(P'_k, Q'_k, n'_k) \in \mathcal{R}$; we have

$$T^+(P_k, Q_k, n_k) = (P'_k, Q'_k, n'_k) \tag{10.56}$$

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

$$g^n(P_k) \subset P'_k, \qquad g^n(\omega) \subset \omega' \tag{10.57}$$

and we also define

$$\widetilde{T}^+ / \widetilde{\mathcal{R}}^\infty_+(P) = g^n / \widetilde{\mathcal{R}}^\infty_+(P). \tag{10.58}$$

We, thus, have a commutative diagram

$$\begin{array}{ccc} \widetilde{\mathcal{D}}_+ & \xrightarrow{\widetilde{T}^+} & \widetilde{\mathcal{R}}^\infty_+ \\ \downarrow \pi & & \downarrow \pi \\ \mathcal{D}_+ & \xrightarrow{T^+} & \mathcal{R}^\infty_+ \end{array} \tag{10.59}$$

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 (10.55); the intersection $\omega = \bigcap_{k\geqslant 0} P_k$ 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

$$\widetilde{T}^+(\omega) = \omega' \cap Q. \tag{10.60}$$

10.5. *Topology and geometry of $\mathcal{R}^\infty_+$ and $\widetilde{\mathcal{R}}^\infty_+$.* — Each stable curve is a compact subset of $R = \bigcup R_a$. Therefore, $\mathcal{R}^\infty_+$ 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 $\mathcal{R}^\infty_+$ can also be viewed directly: for any $\omega = \bigcap P_k$ in $\mathcal{R}^\infty_+$, 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, $\mathcal{R}^\infty_+$ 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 $C^{1+\mathrm{Lip}}$ map such that $\omega = \{x_a = \varphi_\omega(y_a)\}$; for each $a \in \mathcal{A}$, each $y^0_a \in I^u_a$, the map

$$\begin{aligned} \phi_{y^0_a} : \mathcal{R}^\infty_+(a) &\mapsto I^s_a, \\ \omega &\mapsto \varphi_\omega(y^0_a) \end{aligned} \tag{10.61}$$

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 $\widetilde{\mathcal{R}}_+^\infty(a)$.

Regarding the transverse regularity of the partial foliation $\widetilde{\mathcal{R}}_+^\infty(a)$, we have the following result.

Proposition **51.** — *For all $a \in \mathfrak{a}$, all distinct $\omega, \omega' \in \mathcal{R}_+^\infty(a)$, all $y \in I_a^u$, we have*

$$\left| \frac{\partial}{\partial y} \log |\varphi_\omega - \varphi_{\omega'}| \right| \leqslant C.$$

In particular, the homeomorphisms $\phi_{y'} \circ \phi_y^{-1}$ are bi-Lipschitzian, uniformly in y, y'.

Proof. — For $(P, Q, n) \in \mathcal{R}$, let $y = \varphi_P^\pm(x)$ be the vertical-like curves bounding P. We will prove that, for $(P, Q, n), (P', Q', n') \in \mathcal{R}$ with P, P' contained in the same R_a and $P \cap P' = \emptyset$, $\varepsilon, \varepsilon' \in \{\pm\}$, we have, for all $y \in I_a^u$

$$\left| \frac{\partial}{\partial y} \log |\varphi_P^\varepsilon - \varphi_{P'}^{\varepsilon'}| \right| \leqslant C. \tag{10.62}$$

This clearly implies the estimate of the proposition. To prove (10.62), we first observe that the inequality is obvious if we allowed the constant in the right-hand term to depend on P, P'. Let $N > 0$ and $C(N)$ be the smallest constant such that (10.62) is satisfied when $(P, Q, n), (P', Q', n') \in \mathcal{R}$ are as required with $n, n' < N$. We will show that $C(N)$ stays bounded.

Let $(P, Q, n), (P', Q', n') \in \mathcal{R}$ with P, P' contained in the same R_a, $P \cap P' = \emptyset$ and $\max(n, n') = N$. Let $(\widetilde{P}, \widetilde{Q}, \widetilde{n}) \in \mathcal{R}$ be the element such that $\widetilde{P}$ is the smallest rectangle containing both P and P'. Let $\overline{P}, \overline{P}'$ be the children of $\widetilde{P}$ which contain P, P' respectively. We distinguish two cases.

1. *At least one of $\overline{P}, \overline{P}'$ is a simple child.*
   In this case, we observe that, for all $y \in I_a^u$, we have

$$C^{-1} |\widetilde{P}| \leqslant |\varphi_P^\varepsilon(y) - \varphi_{P'}^{\varepsilon'}(y)| \leqslant C |\widetilde{P}|. \tag{10.63}$$

On the other hand, we have seen in the proof of Proposition 50 (cf. (10.52)) that

$$\begin{aligned} |D\varphi_P^\varepsilon(y) - D\varphi_{\widetilde{P}}^+(y)| &\leqslant C |\widetilde{P}|, \\ |D\varphi_{P'}^{\varepsilon'}(y) - D\varphi_{\widetilde{P}}^+(y)| &\leqslant C |\widetilde{P}|. \end{aligned} \tag{10.64}$$

We obtain (10.62) with some uniform constant C_0.

2. *Both* $\overline{P}, \overline{P}'$ *are non-simple children.*
Let $\overline{\varepsilon}, \overline{\varepsilon}'$ such that

$$|\varphi_P^{\varepsilon}(y) - \varphi_{P'}^{\varepsilon'}(y)| \geqslant |\varphi_{\overline{P}}^{\overline{\varepsilon}}(y) - \varphi_{\overline{P}'}^{\overline{\varepsilon}'}(y)|.$$

We apply the proposition in Appendix B, whose hypotheses are clearly satisfied. By definition of C(N), we obtain

$$\left|\frac{\partial}{\partial y}\log|\varphi_{\overline{P}}^{\overline{\varepsilon}} - \varphi_{\overline{P}'}^{\overline{\varepsilon}'}|\right| \leqslant C(C(N)|\widetilde{Q}\rfloor^{1/2} + 1). \tag{10.65}$$

From (10.52), we have

$$|D\varphi_P^{\varepsilon}(y) - D\varphi_{\overline{P}}^{\overline{\varepsilon}}(y)| \leqslant C|\overline{P}|,$$

$$|D\varphi_{P'}^{\varepsilon'}(y) - D\varphi_{\overline{P}'}^{\overline{\varepsilon}'}(y)| \leqslant C|\overline{P}'|$$

and from (MP7)

$$|\varphi_P^{\varepsilon}(y) - \varphi_{P'}^{\varepsilon'}(y)| \geqslant C^{-1}(|\overline{P}| + |\overline{P}'|).$$

In view of this and (10.65), (10.66), we obtain (10.62) with a constant $C = C_1(1 + |\widetilde{Q}\rfloor^{1/2}C(N))$ for some uniform constant C_1.

Observe that, when the second case occurs, N has to be large (at least of the order of $\log\varepsilon_0^{-1}$) and we must have $|\widetilde{Q}\rfloor \ll \varepsilon_0$. Therefore, one has $C(N) \leqslant C_0$ for $N \ll \log\varepsilon_0^{-1}$ and

$$C(N+1) \leqslant \max(C_0, C_1(1 + C(N)\varepsilon_0^{1/2})). \tag{10.66}$$

This implies that C(N) is bounded (by $\max(C_0, 2C_1)$, as soon as $\varepsilon_0^{1/2}\max(C_0, 2C_1) < 1$) and ends the proof of the proposition. □

The result of Proposition 51 implies that the transverse Hausdorff dimension $d_s = d_s(g)$ of $\widetilde{\mathcal{R}}_+^{\infty}$ 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 $\widetilde{R}_+^{\infty} \cap P_{aa'}$ into $\widetilde{R}_+^{\infty} \cap R_{a'}$; the transverse Hausdorff dimension of $\widetilde{R}_+^{\infty} \cap R_{a'}$ is therefore not smaller than that of $\widetilde{R}_+^{\infty} \cap R_a$; 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.

Let $\omega, \widehat{\omega}$ be two stable curves and $j > 0$. Assume that ω and $\widehat{\omega}$ belong to the domain of $(T^+)^j$. We say that ω and $\widehat{\omega}$ belong to the same component of the domain of $(T^+)^j$ if, for each $0 \leqslant i < j$, there exists a prime P_i such that $(T^+)^i(\omega)$ and $(T^+)^i(\widehat{\omega})$ belong to $\mathcal{R}_+^{\infty}(P_i)$.

Proposition **52.** — *There exists* $\theta_0 \in (0,1)$ *such that, if* ω, $\widehat{\omega}$ *belong to the same component of the domain of* $(T^+)^j$, *we have*

$$|\varphi_\omega(y) - \varphi_{\widehat{\omega}}(y)| \leqslant C\theta_0^j,$$
$$|D\varphi_\omega(y) - D\varphi_{\widehat{\omega}}(y)| \leqslant C\theta_0^j,$$

for all y.

Proof. — With $(T^+)^i(\omega) \in \mathcal{R}_+^\infty(P_i)$ for $0 \leqslant i < j$, let $(P, Q, n) = (P_0, Q_0, n_0) * \cdots * (P_{j-1}, Q_{j-1}, n_{j-1})$. We have $\omega, \widehat{\omega} \subset P$ and $|P| \leqslant C\theta_0^j$ for some fixed $\theta_0 \in (0,1)$. The estimates are therefore a consequence of those in Proposition 50. □

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

$$v_\omega(z) = \frac{\partial}{\partial y} + D\varphi_\omega(y)\frac{\partial}{\partial x} \tag{10.67}$$

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

$$A_x^{-1}(y, x')\left(1 - B_x(y, x')\, D\varphi_{\omega'}(y')\right). \tag{10.68}$$

We denote by $\widetilde{b}(z)$ the logarithm of the absolute value of this coefficient. As φ_ω is $C^{1+\mathrm{Lip}}$ uniformly in ω and $g^n : P \mapsto Q$ has bounded distortion, we have, for all $z, z^* \in \omega$

$$|\widetilde{b}(z) - \widetilde{b}(z^*)| \leqslant C|z - z^*|. \tag{10.69}$$

Let $j > 0$ and let ω, $\widehat{\omega}$ be stable curves which belong to the same component of the domain of $(T^+)^j$. Let z, $\hat{z}$ be point of ω, $\widehat{\omega}$, respectively, with the same y coordinate. It follows from Proposition 52 that one has

$$|\widetilde{b}(z) - \widetilde{b}(\hat{z})| \leqslant C\theta_0^j. \tag{10.70}$$

We also have, from the definition of $\widetilde{b}$:

$$|\widetilde{b}(z) + \log|P|| \leqslant C. \tag{10.71}$$

We want to get rid of the dependence of $\widetilde{b}$ on the y coordinate along ω by adding a coboundary term. We define

$$\mathcal{D}_+^\infty = \bigcap_{j\geqslant 0} \mathrm{Dom}(\mathrm{T}^+)^j = \bigcap_{j\geqslant 0} (\mathrm{T}^+)^{-j}(\mathcal{D}_+). \tag{10.72}$$

For each a, fix some $y_a^0 \in \mathrm{I}_a^u$. Then, for $\omega \in \mathcal{D}_+^\infty$, $\omega \subset \mathrm{R}_a$, $z \in \omega$, define

$$\Delta b(z) = \sum_{i\geqslant 0} \Big(\widetilde{b}((\widetilde{\mathrm{T}}^+)^i(z)) - \widetilde{b}((\widetilde{\mathrm{T}}^+)^i(z^0)) \Big), \tag{10.73}$$

where z^0 is the point on ω with vertical coordinate equal to y_a^0.

From the cone condition, we have, for $i \geqslant 0$:

$$\|(\widetilde{\mathrm{T}}^+)^i(z) - (\widetilde{\mathrm{T}}^+)^i(z^0)\| \leqslant \mathrm{C}\lambda^{-i}. \tag{10.74}$$

The series defining Δb is uniformly convergent from (10.69), (10.74), and Δb is bounded on $\widetilde{\mathcal{D}}_+^\infty := \pi^{-1}(\mathcal{D}_+^\infty)$.

Write z^1 for the point on $\mathrm{T}_+(\omega) \in \mathrm{R}_{a'}$ with vertical coordinate $y_{a'}^0$. We have

$$\Delta b(z) - \Delta b(\widetilde{\mathrm{T}}^+(z)) = \widetilde{b}(z) - b(\omega), \tag{10.75}$$

with

$$b(\omega) = \widetilde{b}(z^0) + \sum_{i\geqslant 0} \Big[\widetilde{b}((\widetilde{\mathrm{T}}^+)^{i+1}(z^0)) - \widetilde{b}((\widetilde{\mathrm{T}}^+)^i(z^1)) \Big]. \tag{10.76}$$

We call b the (logarithmic) mean transverse dilatation.

Proposition **53.** — *The mean transverse dilatation b, which differs from $\widetilde{b}$ on $\mathcal{D}_+^\infty$ by the coboundary of the bounded function Δb, satisfies*

$$|b(\omega) - b(\widehat{\omega})| \leqslant \mathrm{C}\theta_1^j \tag{10.77}$$

if ω, $\widehat{\omega}$ belong to the same component of the domain of $(\mathrm{T}^+)^j$. Here θ_1 is a fixed constant in $(0,1)$ larger than θ_0.

Proof. — We have only to prove (10.77). Let $z^0 \in \omega$, $z^1 \in \mathrm{T}^+(\omega)$ as above and let $\hat{z}^0 \in \widehat{\omega}$, $\hat{z}^1 \in \mathrm{T}^+(\widehat{\omega})$ be similarly defined. We have, for $i \geqslant 0$, using (10.69) and (10.74):

$$|\widetilde{b}((\widetilde{\mathrm{T}}^+)^{i+1}(z^0)) - \widetilde{b}((\widetilde{\mathrm{T}}^+)^i(z^1))| \leqslant \mathrm{C}\lambda^{-i}, \tag{10.78}$$

$$|\widetilde{b}((\widetilde{\mathrm{T}}^+)^{i+1}(\hat{z}^0)) - \widetilde{b}((\widetilde{\mathrm{T}}^+)^i(\hat{z}^1))| \leqslant \mathrm{C}\lambda^{-i}. \tag{10.79}$$

From (10.70), we also have

$$|\widetilde{b}(z^0) - \widetilde{b}(\hat{z}^0)| \leqslant \mathrm{C}\theta_0^j. \tag{10.80}$$

For $0 \leqslant i < j/2$, we will compare $\widetilde{b}((\widetilde{T}^+)^{i+1}(z^0))$ and $\widetilde{b}((\widetilde{T}^+)^{i+1}(\hat{z}^0))$ as follows. Let $\widetilde{z}^{i+1}$ be the point on $T_+^{i+1}(\widehat{\omega})$ with the same y-coordinate as $(\widetilde{T}^+)^{i+1}(z^0)$. From (10.70), we have

$$|\widetilde{b}(\widetilde{z}^{i+1}) - \widetilde{b}((\widetilde{T}^+)^{i+1}(z^0))| \leqslant C\theta_0^{j-i-1}, \tag{10.81}$$

From Proposition 52, the cone condition, and (10.69), we also have

$$\begin{aligned} \|\widetilde{z}^{i+1} - (\widetilde{T}^+)^{i+1}(z^0)\| &\leqslant C\theta_0^{j-i-1}, \\ \|\widetilde{z}^{i+1} - (\widetilde{T}^+)^{i+1}(\hat{z}^0)\| &\leqslant C\theta_0^{j-i-1}, \\ |\widetilde{b}(\widetilde{z}^{i+1}) - \widetilde{b}((\widetilde{T}^+)^{i+1}(\hat{z}^0))| &\leqslant C\theta_0^{j-i-1}. \end{aligned} \tag{10.82}$$

We therefore have

$$|\widetilde{b}((\widetilde{T}^+)^{i+1}(\hat{z}^0))) - \widetilde{b}((\widetilde{T}^+)^{i+1}(z^0)))| \leqslant C\theta_0^{j-i-1}. \tag{10.83}$$

Similarly, we get, for $0 \leqslant i < j/2$

$$|\widetilde{b}((\widetilde{T}^+)^{i}(\hat{z}^1))) - \widetilde{b}((\widetilde{T}^+)^{i}(z^1)))| \leqslant C\theta_0^{j-i}. \tag{10.84}$$

To estimate the difference $b(\omega) - b(\widehat{\omega})$, we use first (10.80), then (10.78), (10.79) to truncate the sums at $j/2$, and finally (10.81), (10.82) to estimate the difference of the remaining terms. We obtain (10.77) with

$$\theta_1 = \max(\theta_0^{1/2}, \lambda^{-1/2}). \tag{10.85}$$

□

10.7. *Definition of a transfer operator.* — As Δb is bounded, it follows from (10.71) that

$$|b(\omega) + \log|P|| \leqslant C \tag{10.86}$$

for all $(P, Q, n) \in \mathcal{P}$, $\omega \in \mathcal{R}_+^\infty(P)$.

It is then a consequence of Corollary 16 in Section 10.3 that the series

$$\sum_{T^+\omega'=\omega} \exp(-db(\omega')) \tag{10.87}$$

over pre-images ω' of a given stable curve ω is converging, uniformly in ω, for $d \geqslant d_s^-$. Here $d_s^- < d_s^t$ should satisfy the hypotheses of Proposition 49 and Corollary 16. We will, therefore, define a transfer operator L_d for $d \geqslant d_s^-$ as follows: for a bounded function h defined on $\mathcal{D}_+^\infty$, for $\omega \in \mathcal{D}_+^\infty$, we set

$$L_d h(\omega) = \sum_{T^+\omega'=\omega} \exp(-db(\omega'))h(\omega'). \tag{10.88}$$

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. Section 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 (10.88) defines a bounded operator from the space of bounded functions on $\mathcal{D}_+^\infty$ 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

$$\theta_1 < \theta < 1 \tag{10.89}$$

where θ_1 comes from Proposition 53 and satisfies $\theta_1 > \lambda^{-1}$ (cf. (10.85)). Denote by E the space of bounded functions h on $\mathcal{D}_+^\infty$ which satisfy, for some constant $C > 0$,

$$|h(\omega) - h(\widehat{\omega})| \leqslant C\theta^j \tag{10.90}$$

whenever ω, $\widehat{\omega}$ 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 in (10.90), and set

$$\|h\|_E = \max(|h|_E, \|h\|_\infty). \tag{10.91}$$

It is clear that E is a Banach space.

Proposition **54.** — *For $d \geqslant d_s^-$, 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 ω, $\widehat{\omega}$ belong to the same component of the domain of $(T^+)^j$. Let (P, Q, n) be a prime such that Q, ω, $\widehat{\omega}$ belong to the same rectangle R_a, and let ω_1, $\widehat{\omega}_1$ be the inverse images of ω, $\widehat{\omega}$ by T^+ corresponding to this inverse branch. By the definition of $|\ |_E$, we have

$$|h(\omega_1) - h(\widehat{\omega}_1)| \leqslant |h|_E \theta^{j+1}. \tag{10.92}$$

From Proposition 53, we have

$$|b(\omega_1) - b(\widehat{\omega}_1)| \leqslant C\theta_1^{j+1}. \tag{10.93}$$

It follows from (10.86) that

$$|\exp(-db(\omega_1)) - \exp(-db(\widehat{\omega}_1))| \leqslant C(d)|P|^d \theta_1^{j+1}. \tag{10.94}$$

Putting together (10.92) and (10.94), we have

$$\begin{aligned} &|h(\omega_1)\exp(-db(\omega_1)) - h(\widehat{\omega}_1)\exp(-db(\widehat{\omega}_1))| \\ &\quad \leqslant C(d)|P|^d(\theta^{j+1}|h|_E + \theta_1^{j+1}\|h\|_\infty). \end{aligned} \tag{10.95}$$

Summing over (non trivial) primes yields for $d \geqslant d_s^-$:

$$(\mathbf{10.96}) \qquad |\Delta \mathrm{L}_d h|_{\mathrm{E}} < \varepsilon_1 \|h\|_{\mathrm{E}},$$

$$(\mathbf{10.97}) \qquad |\mathrm{L}_d h|_{\mathrm{E}} < \mathrm{C}\|h\|_{\mathrm{E}},$$

where ε_1 can be made arbitrarily small if ε_0 is small enough, according to Corollary 16. The same estimates (for $d \geqslant d_s^-$) for $\|\Delta \mathrm{L}_d h\|_\infty$ and $\|\mathrm{L}_d h\|_\infty$ are easier and can be seen directly. The proposition follows. □

10.8. *Spectral properties of the transfer operator.* — Let us denote by $\mathcal{R}_+^\infty(\mathrm{K})$ the set of stable curves ω which are intersections of a sequence of rectangles belonging to $\mathcal{R}(\mathrm{I}_0)$; these stable curves are precisely those which meet the initial horseshoe K.

Observe that $\mathcal{R}_+^\infty(\mathrm{K}) \subset \mathcal{D}_+^\infty$. Denote by $\mathrm{E_K}$ the space of bounded functions h on $\mathcal{R}_+^\infty(\mathrm{K})$ which satisfy

$$(\mathbf{10.98}) \qquad |h(\omega) - h(\widehat{\omega})| \leqslant \mathrm{C}\theta^j,$$

whenever ω, $\widehat{\omega}$ belong to the same component of the domain of $(\mathrm{T}^+)^j$. Define $|h|_{\mathrm{E_K}}$, $\|h\|_{\mathrm{E_K}}$ as above, which makes $\mathrm{E_K}$ a Banach space.

Let $h \in \mathrm{E}$; the restriction of h to $\mathcal{R}_+^\infty(\mathrm{K})$ belongs to $\mathrm{E_K}$ and we have

$$(\mathbf{10.99}) \qquad \|h/\mathcal{R}_+^\infty(\mathrm{K})\|_{\mathrm{E_K}} \leqslant \|h\|_{\mathrm{E}}.$$

The formula for L_d^0 defines a bounded operator, still denoted by L_d^0, on $\mathrm{E_K}$ and we have a commutative diagram

$$(\mathbf{10.100}) \qquad \begin{array}{ccc} \mathrm{E} & \xrightarrow[\mathrm{L}_d^0]{} & \mathrm{E} \\ {\scriptstyle r}\downarrow & & \downarrow{\scriptstyle r} \\ \mathrm{E_K} & \xrightarrow[\mathrm{L}_d^0]{} & \mathrm{E_K} \end{array}$$

where $r : \mathrm{E} \mapsto \mathrm{E_K}$ is the restriction operator. The bounded operator $\mathrm{L}_d^0 : \mathrm{E_K} \mapsto \mathrm{E_K}$ 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

$$(\mathbf{10.101}) \qquad \mathrm{E_K} = \mathbf{R}h'_d \oplus \mathrm{H}'_d$$

depending analytically on the parameter d, such that h'_d is a positive eigenfunction, with associated eigenvalue $\lambda'_d > 0$, and such that

$$(\mathbf{10.102}) \qquad sp\left(\mathrm{L}_d^0/\mathrm{H}'_d\right) \subset \{|z| < \lambda'_d\}.$$

(b) There exists a (unique) probability measure μ'_d on $\mathcal{R}^\infty_+(\mathrm{K})$ such that

$$\mathrm{H}'_d = \left\{h \in \mathrm{E_K}, \int h d\mu'_d = 0\right\}. \tag{10.103}$$

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 $\mathcal{R}^\infty_+(\mathrm{K})$ (observe that $\widetilde{\mathrm{T}}^+$ on $\widetilde{\mathrm{R}}^\infty_+(\mathrm{K})$ is just the restriction of g).

Let E^0 be the kernel of the restriction operator $r : \mathrm{E} \mapsto \mathrm{E_K}$. It is invariant under L^0_d.

Lemma **17.** — *One has, for all* $d \in \mathbf{R}$.

$$sp\left(\mathrm{L}^0_d/\mathrm{E}^0\right) \subset \{|z| \leqslant \theta\lambda'_d\}.$$

Proof. — Let $h \in \mathrm{E}^0$, $j \geqslant 0$. We have

$$(\mathrm{L}^0_d)^j h(\omega) = \sum^0_{(\mathrm{T}^+)^j(\omega')=\omega} h(\omega') \exp(-db^{(j)}(\omega')), \tag{10.104}$$

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

$$b^{(j)}(\omega') = \sum_{0 \leqslant i < j} b((\mathrm{T}^+)^i(\omega')). \tag{10.105}$$

We observe that in the sum in (10.104), each ω' belongs to the same component of the domain of $(\mathrm{T}^+)^j$ as a stable curve in $\mathcal{R}^\infty_+(\mathrm{K})$. As h belongs to E^0, this implies that for such a ω' we have

$$|h(\omega')| \leqslant |h|_{\mathrm{E}} \theta^j. \tag{10.106}$$

On the other hand, we have

$$\sum^0 \exp(-db^{(j)}(\omega')) \leqslant \mathrm{C}\lambda'^j_d, \tag{10.107}$$

and it follows that

$$\|(\mathrm{L}^0_d)^j h\|_\infty \leqslant \mathrm{C}\lambda'^j_d \theta^j \|h\|_{\mathrm{E}}. \tag{10.108}$$

Let $\widehat{\omega} \in \mathcal{R}^\infty_+$ belong to the same component of the domain of $(\mathrm{T}^+)^\ell$ as ω. Denote by $\widehat{\omega}'$ the inverse image of $\widehat{\omega}$ associated to the same sequence of trivial primes as ω'. We have

$$|h(\omega') - h(\widehat{\omega}')| \leqslant |h|_{\mathrm{E}} \theta^{j+\ell}, \tag{10.109}$$

and, from Proposition 53

$$(\mathbf{10.110})\qquad |b^{(j)}(\omega') - b^{(j)}(\widehat{\omega}')| \leqslant C\theta_1^{\ell}.$$

Using also (10.106) and (10.107), we obtain

$$(\mathbf{10.111})\qquad \left|(L_d^0)^j h(\omega) - (L_d^0)^j h(\widehat{\omega})\right| \leqslant C\theta^{j+\ell}\lambda_d'^j \|h\|_E,$$

which implies the statement of the Lemma. □

We deduce from Lemma 17 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

$$(\mathbf{10.112})\qquad L_d^0(h'_d) = \lambda'_d h'_d.$$

Moreover, defining a supplementary hyperplane by

$$(\mathbf{10.113})\qquad H''_d = r^{-1}(H'_d) \oplus E^0,$$

we have that H''_d is invariant under L_d^0 and

$$(\mathbf{10.114})\qquad sp\Big(L_d^0/H''_d\Big) \subset \{|z| \leqslant \lambda''_d\},$$

where $\lambda''_d < \lambda'_d$ is independent of ε_0.

Using Proposition 54, we now consider L_d itself, assuming that ε_0 is small enough and $d \geqslant d_s^-$.

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

$$(\mathbf{10.115})\qquad sp\Big(L_d/H_d\Big) \subset \{|z| < \lambda_d\}.$$

Moreover, h_d, λ_d and H_d depend analytically on d for $d > d_s^-$ because L_d does. We check that

$$(\mathbf{10.116})\qquad h_d \geqslant C^{-1} > 0.$$

Indeed, the sequence $h^{(n)} = \lambda_d^{-n} L_d^n(1)$ converges to a positive multiple of h_d. We have

$$(\mathbf{10.117})\qquad h^{(n)}(\omega) = \lambda_d^{-n} \sum_{(T^+)^n(\omega')=\omega} \exp(-d b^{(n)}(\omega')).$$

Let ω, $\widehat{\omega}$ be elements of $\mathcal{D}_+^\infty$ in the same rectangle R_a; let ω', $\widehat{\omega}'$ be pre-images of ω, $\widehat{\omega}$ by $(T^+)^n$ associated with the same sequence of primes. We have (cf. (10.110))

$$(\mathbf{10.118})\qquad |b^{(n)}(\omega') - b^{(n)}(\widehat{\omega}')| \leqslant C,$$

and it follows that

$$C^{-1} \leqslant (h^{(n)}(\omega'))^{-1} h^{(n)}(\omega) \leqslant C. \tag{10.119}$$

This implies (10.116). One normalizes h_d in order to have

$$h_d = \lim_{n \to +\infty} \lambda_d^{-n} L_d^n(1). \tag{10.120}$$

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 $h \in E$

$$\lim_{n \to \infty} \lambda_d^{-n} L_d^n h = \mu_d(h) h_d. \tag{10.121}$$

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 $\mathcal{R}_+^\infty$, 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 (10.121) of μ_d, we have, for all $h \in E$

$$\mu_d(L_d h) = \lambda_d \mu_d(h). \tag{10.122}$$

We will now check the classical Jacobian property for μ_d.

Let $(P, Q, n) \in \mathcal{P}$, with $Q \subset R_a$. The application T^+ is a bijection T_P from the set $\mathcal{R}_+^\infty(P)$ onto $\mathcal{R}_+^\infty(a)$.

Let h be a function in E which vanishes outside $\mathcal{R}_+^\infty(P)$. Then, $L_d h$ vanishes outside $\mathcal{R}_+^\infty(a)$, and satisfies on $\mathcal{R}_+^\infty(a)$

$$L_d h(\omega) = h(T_P^{-1}\omega) \exp(-db(T_P^{-1}\omega)).$$

Plugging this into (10.122) gives

$$\lambda_d \int h(\omega) d\mu_d(\omega) = \int_{\mathcal{R}_+^\infty(a)} h(T_P^{-1}\omega) \exp(-db(T_P^{-1}\omega)) d\mu_d(\omega). \tag{10.123}$$

This relation is the Jacobian property of the measure μ_d.

Consider in particular the case where h is the characteristic function of $\mathcal{R}_+^\infty(P)$. We then obtain

$$\lambda_d \mu_d(\mathcal{R}_+^\infty(P)) = \int_{\mathcal{R}_+^\infty(a)} \exp(-db(T_P^{-1}\omega)) d\mu_d(\omega). \tag{10.124}$$

We now will specify the value of d by asking that

$$\lambda_d = 1. \tag{10.125}$$

Indeed, we have the following

Proposition **55.** — *One has* $\frac{\partial}{\partial d}\lambda_d < 0$ *for* $d \geqslant d_s^-$, *and also* $\lambda_{d_s^-} > 1$, $\lim_{d\to+\infty}\lambda_d = 0$. *Therefore, there exists a unique* $d \geqslant d_s^-$ *with* $\lambda_d = 1$.

Proof. — We first prove that $\frac{\partial}{\partial d}\lambda_d < 0$. The Birkhoff sums of $\widetilde{b}$ for $\widetilde{\mathrm{T}}^+$ grow at least linearly, and $b - \widetilde{b}$ is the coboundary of a bounded function. Therefore, there exists $j > 0$ such that the Birkhoff sum $b^{(j)}$ of b for T^+ is everywhere > 1. For $h \in \mathrm{E}$, we have

$$(\mathbf{10.126}) \qquad \frac{\partial}{\partial d}\mathrm{L}_d^j(h)(\omega) = -\sum_{(\mathrm{T}^+)^j(\omega')=\omega} h(\omega')b^{(j)}(\omega')\exp(-db^{(j)}(\omega'))$$

which is everywhere < 0 is $h > 0$. We will apply this with $h = h_d$. We differentiate with respect to d the relation $\mathrm{L}_d^j(h_d) = \lambda_d^j h_d$ to get

$$(\mathbf{10.127}) \qquad \frac{\partial}{\partial d}\mathrm{L}_d^j(h_d) - \left(\frac{\partial}{\partial d}\lambda_d^j\right)h_d = (\mathrm{L}_d^j - \lambda_d^j)\left(\frac{\partial}{\partial d}h_d\right).$$

The right-hand term belongs to H_d. Applying μ_d shows that $\frac{\partial}{\partial d}\lambda_d^j$ and then also $\frac{\partial}{\partial d}\lambda_d$ is < 0.

Next, d_s^- was chosen in order to be smaller than the transverse Hausdorff dimension of $\mathrm{W}^s(\mathrm{K})$. This means that the eigenvalue $\lambda'_{d_s^-}$ for $\mathrm{L}^0_{d_s^-}$ on $\mathrm{E_K}$ satisfies $\lambda'_{d_s^-} > 1$. As $\Delta\mathrm{L}_{d_s^-}$ is also a nonnegative operator, we have $\lambda_{d_s^-} \geqslant \lambda'_{d_s^-} > 1$.

Finally, with j as above and $h \in \mathrm{E}$, $h > 0$ we have

$$(\mathbf{10.128}) \qquad \mathrm{L}_d^j(h)(\omega) = \sum_{(\mathrm{T}^+)^j(\omega')=\omega} h(\omega')\exp(-db^{(j)}(\omega')) \leqslant e^{d_s^- - d}\mathrm{L}^j_{d_s^-}(h)(\omega)$$

which implies that $\lambda_d \leqslant e^{\frac{1}{j}(d_s^- - d)}\lambda_{d_s^-}$. □

We will denote by d_s the value of d such that $\lambda_d = 1$. We shall indeed see that d_s is the transverse Hausdorff dimension of $\widetilde{\mathcal{R}}_+^\infty$ which we were able to define in Section 10.5.

We just write μ for the measure μ_{d_s} and h^* for the eigenfunction h_{d_s}.

Proposition **56.** — *For any* $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}$, *we have*

$$\mathrm{C}^{-1}|\mathrm{P}|^{d_s} \leqslant \mu(\{\omega \subset \mathrm{P}\}) \leqslant \mathrm{C}|\mathrm{P}|^{d_s}.$$

Proof. — Let

$$(\mathbf{10.129}) \qquad (\mathrm{P}, \mathrm{Q}, n) = (\mathrm{P}_1, \mathrm{Q}_1, n_1) * \cdots * (\mathrm{P}_r, \mathrm{Q}_r, n_r)$$

be the prime decomposition of $(\mathrm{P}, \mathrm{Q}, n)$. If $\omega \in \mathcal{D}_+^\infty$ satisfies $(\mathrm{T}^+)^i(\omega) \in \mathcal{R}_+^\infty(\mathrm{P}_{i+1})$ for $0 \leqslant i < r$, we claim that

$$(\mathbf{10.130}) \qquad \mathrm{C}^{-1}|\mathrm{P}|^{d_s} \leqslant \exp(-d_s b^{(r)}(\omega)) \leqslant \mathrm{C}|\mathrm{P}|^{d_s}$$

(see the definition of $b^{(r)}$ in (10.105)). Indeed, let $z \in \omega$; denoting, by $\widetilde{b}^{(r)}$ the Birkhoff sum of $\widetilde{b}$ for $\widetilde{T}^{+}$, $\widetilde{b}^{(r)}(z)$ is the logarithm of the absolute value of the first diagonal coefficient of the matrix of $D(\widetilde{T}^{+})^{r}$ at z, hence we have, by bounded distortion

$$C^{-1}|P| \leqslant \exp(-\widetilde{b}^{(r)}(z)) \leqslant C|P|. \tag{10.131}$$

On the other hand, as $b - \widetilde{b}$ is the coboundary of a bounded function the difference $|\widetilde{b}^{(r)}(z) - b^{(r)}(\omega)|$ is bounded by C and the claim follows.

From the Jacobian property, we have

$$\begin{aligned} \mu(\{\omega \subset P\}) &\geqslant C^{-1}|P|^{d_s}\mu(\{\omega \in \mathcal{R}_{+}^{\infty}(a)\}) \\ &\geqslant C^{-1}|P|^{d_s}, \end{aligned} \tag{10.132}$$

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 $0 \leqslant i \leqslant r$, let

$$(P^{i}, Q^{i}, n^{i}) = (P_{i+1}, Q_{i+1}, n_{i+1}) * \cdots * (P_r, Q_r, n_r) \tag{10.133}$$

(with $(P^{r}, Q^{r}, n^{r}) = (R_a, R_a, 0)$). We have

$$L_{d_s}\chi_P = \chi_P^1 + \Delta\chi_P^1 \tag{10.134}$$

where

$$\chi_P^1(\omega^1) = \begin{cases} 0 & \text{if } \omega^1 \not\subset P^1, \\ \exp(-d_s b(\omega^0)) & \text{if } \omega^1 = T^{+}(\omega^0) \text{ for some } \omega^0 \in \mathcal{R}_{+}^{\infty}(P) \end{cases} \tag{10.135}$$

and

$$\Delta\chi_P^1 \leqslant C\sum |P_1^*|^{d_s}, \tag{10.136}$$

where the sum runs over prime elements (P_1^*, Q_1^*, n_1^*) with P_1^* contained in P and distinct from P (when $r = 1$). By Proposition 49 in Section 10.3, we obtain

$$\mu(\Delta\chi_P^1) \leqslant C|P|^{d_s}\kappa^{\frac{r-1}{2}}. \tag{10.137}$$

If $r > 1$, we write similarly

$$L_d\chi_P^1 = \chi_P^2 + \Delta\chi_P^2, \tag{10.138}$$

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

$$\Delta\chi_P^2 \leqslant C|P_1|^{d_s}\sum |P_2^*|^{d_s}, \tag{10.139}$$

where the sum now is over primes P_2^* contained in P^1 and distinct from P^1 (when $r=2$). Proposition 49 now gives

$$\begin{aligned}\mu(\Delta\chi_P^2) &\leqslant C|P_1|^{d_s}|P^1|^{d_s}\kappa^{\frac{r-2}{2}}\\ &\leqslant C|P|^{d_s}\kappa^{\frac{r-2}{2}}.\end{aligned} \tag{10.140}$$

We iterate this process. At step i, we will have

$$\mu(\Delta\chi_P^i) \leqslant C|P|^{d_s}\kappa^{\frac{r-i}{2}} \tag{10.141}$$

where the constant C does not gets worse by the same argument used above to justify (10.130).

At the last step, we have from (10.130)

$$\mu(\chi_P^r) \leqslant C|P|^{d_s}. \tag{10.142}$$

The contribution of the perturbative terms is bounded by

$$\mu\left(\sum_1^{r-1}\Delta\chi_P^i\right) \leqslant C|P|^{d_s}\sum_1^r \kappa^{\frac{r-i}{2}} \leqslant C|P|^{d_s}. \tag{10.143}$$

□

Corollary **17.** — *The transverse Hausdorff dimension of $\widetilde{\mathcal{R}}_+^\infty$ is $\leqslant d_s$. More precisely, for any* C^1 *curve γ which is transverse to $\widetilde{\mathcal{R}}_+^\infty$, the Hausdorff measure in dimension d_s of the intersection of γ with $\widetilde{\mathcal{R}}_+^\infty$ 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 $|P_i| \leqslant \delta$ for each i and $\widetilde{\mathcal{R}}_+^\infty \subset \cup P_i$. We have

$$\begin{aligned}1 = \sum \mu(P_i) &\geqslant C^{-1}\sum |P_i|^{d_s}\\ &\geqslant C^{-1}\sum [\operatorname{diam}(\gamma\cap P_i)]^{d_s}\end{aligned} \tag{10.144}$$

and the statement of the Corollary follows. □

The following statement shows that the dynamics T^+ is only undefined on a small set.

Proposition **57.** — *The transverse Hausdorff dimension of the set $\widetilde{\mathcal{R}}_+^\infty - \widetilde{\mathcal{D}}_+^\infty$ is $\leqslant d_s^- < d_s$. Moreover, we have*

$$\mu(\widetilde{\mathcal{R}}_+^\infty - \widetilde{\mathcal{D}}_+^\infty) = 0.$$

Proof. — We have

$$\widetilde{\mathcal{R}}_+^\infty - \widetilde{\mathcal{D}}_+^\infty = \bigcup_{n\geqslant 0} (\mathrm{T}^+)^{-n}(\mathcal{N}_+). \tag{10.145}$$

As each $(\mathrm{T}^+)^n$ has countably many inverse branches which are Lipschitzian, it is sufficient to prove that the transverse Hausdorff dimension of $\mathcal{N}_+$ is $\leqslant d_s^-$ and that $\mu(\mathcal{N}_+) = 0$. By the definition of $\mathcal{N}_+$, for any $\delta > 0$, the union of prime rectangles P with $|\mathrm{P}| < \delta$ contains $\mathcal{N}_+$. It then follows from Corollary 16 (in Section 10.3) that the Hausdorff dimension of $\mathcal{N}_+$ is $\leqslant d_s^-$ and (using also Proposition 56) that $\mu(\mathcal{N}_+) = 0$. □

10.10. *Transverse Hausdorff dimension of* $\widetilde{\mathcal{R}}_+^\infty$.

Theorem **4.** — *The transverse Hausdorff dimension of* $\widetilde{\mathcal{R}}_+^\infty$ *is the number* d_s *characterized by* $\lambda_{d_s} = 1$.

Remark **17.** — 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.

10.10.1. *Proof.* — Let γ be a horizontal segment in some R_a. We denote by $[\gamma]$ the set of stable curves which meet γ. We will show that, for all γ, we have

$$\mathrm{A}(\gamma) := \frac{\mu([\gamma])}{(\operatorname{diam}\gamma)^{d_s}} \leqslant \mathrm{C}\log^{\mathrm{C}_0}(\operatorname{diam}\gamma)^{-1}. \tag{10.146}$$

This clearly implies that the transverse Hausdorff dimension of $\widetilde{\mathcal{R}}_+^\infty$ is $\geqslant d_s$, which is sufficient to prove the theorem in view of Corollary 17.

10.10.2. We start with some preliminary work. In the rectangle R_{a_s} which contains L_s, we choose a horizontal line $\mathcal{J} = \{y_s = y_s^*\}$. We will use $\bar{x}$ to denote the x_s coordinate on $\mathcal{J}$ (we use a different notation because we will have in the same formulas points on $\mathcal{J}$ and points in R_{a_s}). Let J be the set of $\bar{x}$ such that $(\bar{x}, y_s^*) \in \mathrm{P}_s$, and J^∞ the set of $\bar{x} \in \mathrm{J}$ such that $(\bar{x}, y_s^*) \in \mathrm{P}_s \cap \widetilde{\mathcal{R}}_+^\infty$. For $\bar{x} \in \mathrm{J}^\infty$, let $x_s = \varphi(y_s, \bar{x})$ be the equation of the stable curve through $(\bar{x}, y_s^*)$ (thus, we have $\bar{x} = \varphi(y_s^*, \bar{x})$).

For each $\bar{x} \in \mathrm{J}^\infty$, φ is a $\mathrm{C}^{1+\mathrm{Lip}}$ function of y_s. Moreover, from (R4) in Section 5.3 and Proposition 50 in Section 10.4, we have

$$|\varphi_y(y_s, \bar{x})| \leqslant \mathrm{C}\varepsilon_0, \tag{10.147}$$

$$|\varphi_y(y_s, \bar{x}) - \varphi_y(y_s', \bar{x})| \leqslant \mathrm{C}\varepsilon_0 |y_s - y_s'|. \tag{10.148}$$

On the other hand, it follows from Proposition 51 in Section 10.5 that, for $\overline{x}, \overline{x}' \in J^\infty$ with $\overline{x}' > \overline{x}$, we have

$$C^{-1}(\overline{x}' - \overline{x}) \leqslant \varphi(y_s, \overline{x}') - \varphi(y_s, \overline{x}) \leqslant C(\overline{x}' - \overline{x}). \tag{10.149}$$

We extend the definition of φ letting $\overline{x}$ run in the whole interval J, to obtain an homeomorphism from $I^u_{a_s} \times J$ onto P_s and still having (10.147)–(10.149), with now $\overline{x}, \overline{x}' \in J$. This can be done for instance by linear interpolation in the $\overline{x}$ variable, for each fixed y_s.

The next step is to switch, via the diffeomorphism G_+ of Section 2.3, from the coordinates x_s, y_s to the coordinates x_s, w. We have (with the notations of Section 2.3) $y_s = Y_s(w, x_s)$; plugging this into φ gives a family of curves parametrized by $\overline{x}$

$$x_s = \varphi(Y_s(w, x_s), \overline{x}). \tag{10.150}$$

Lemma **18.** — *Each curve* $\{x_s = \varphi(Y_s(w, x_s), \overline{x})\}$ *is a graph* $\{x_s = \phi(w, \overline{x})\}$ *of a* $C^{1+\mathrm{Lip}}$ *function of* w. *The function* ϕ *satisfies the same relations* (10.147)–(10.149) *than* φ, *namely*

$$|\phi_w(w, \overline{x})| \leqslant C\varepsilon_0,$$
$$|\phi_w(w, \overline{x}) - \phi_w(w', \overline{x})| \leqslant C\varepsilon_0 |w - w'|,$$
$$C^{-1}(\overline{x}' - \overline{x}) \leqslant \phi(w, \overline{x}') - \phi(w, \overline{x}) \leqslant C(\overline{x}' - \overline{x}),$$

for all $w, w', \overline{x} < \overline{x}'$.

Proof. — In view of (10.147), the first statement follows from the implicit function theorem, which gives also

$$\phi_w = \varphi_y Y_{s,w}(1 - \varphi_y Y_{s,x})^{-1}. \tag{10.151}$$

The first two estimates of the lemma now follow from (10.147), (10.148) and the fact that the partial derivatives of Y_s of first and second order are bounded.

For the last inequality, let $x_s = \phi(w, \overline{x})$, $x'_s = \phi(w, \overline{x}')$. We have $x_s = \varphi(Y_s(w, x_s), \overline{x})$, $x'_s = \varphi(Y_s(w, x'_s), \overline{x}')$ and let $x^*_s = \varphi(Y_s(w, x_s), \overline{x}')$. Then, from (10.149), we get

$$C^{-1}(\overline{x}' - \overline{x}) \leqslant x^*_s - x_s \leqslant C(\overline{x}' - \overline{x}). \tag{10.152}$$

From (10.147), one obtains

$$|x'_s - x^*_s| \leqslant C\varepsilon_0(x'_s - x_s), \tag{10.153}$$

from which we deduce as required

$$C^{-1}(\overline{x}' - \overline{x}) \leqslant x'_s - x_s \leqslant C(\overline{x}' - \overline{x}). \tag{10.154}$$

□

10.10.3. In the rectangle R_{a_u} which contains L_u, consider a C^2 curve $\gamma = \{y_u = \psi(x_u)\}$ contained in Q_u, where ψ satisfies

$$|\psi_x| \leqslant C\varepsilon_0, \qquad |\psi_{xx}| \leqslant C\varepsilon_0. \tag{10.155}$$

Using the diffeomorphism G_- of Section 2.3 to switch to the coordinates w, y_u via $x_u = X_u(w, y_u)$, the curve is transformed into $\{y_u = \psi(X_u(w, y_u))\}$. By the implicit function theorem, this is still a graph $\{y_u = \Psi(w)\}$, with Ψ satisfying

$$|\Psi_w| \leqslant C\varepsilon_0, \qquad |\Psi_{ww}| \leqslant C\varepsilon_0. \tag{10.156}$$

In the spirit of Section 3.5, we now introduce

$$C(w, \bar{x}) := w^2 - \theta(\Psi(w), \phi(w, \bar{x})). \tag{10.157}$$

Observe that, for each $\bar{x}$, the zeros of $C(w, \bar{x})$ correspond to the points of intersection of the curve $G(\gamma \cap L_u)$ with the curve $x_s = \varphi(y_s, \bar{x})$.

Lemma **19.**

1. *For each* $\bar{x} \in J$, $C(w, \bar{x})$ *is a* C^{1+Lip} *function of* w, *satisfying*

$$|C_w(w, \bar{x}) - 2w| \leqslant C\varepsilon_0,$$
$$|C_w(w, \bar{x}) - C_w(w', \bar{x}) - 2(w - w')| \leqslant C\varepsilon_0 |w - w'|.$$

2. *For each* $\bar{x} \in J$, $C(w, \bar{x})$ *attains its minimum value at a unique point* w^*, *which is in the interior of the domain of definition of the* w *variable. Writing*

$$\delta(\bar{x}) := -\min_w C(w, \bar{x}),$$

one has, for $\bar{x} < \bar{x}'$ *in* $\mathcal{J}$

$$C^{-1}(\bar{x}' - \bar{x}) \leqslant |\delta(\bar{x}') - \delta(\bar{x})| \leqslant C(\bar{x}' - \bar{x}).$$

3. *Let* $\bar{x}_0, \bar{x}_1 \in J$ *such that* $\delta(\bar{x}_0) > \delta(\bar{x}_1) > 0$ *and let* w_0, w_1 *such that* $C(w_0, \bar{x}_0) = C(w_1, \bar{x}_1) = 0$. *We have*

$$\delta(\bar{x}_0)^{1/2} |w_0 - w_1| \geqslant C^{-1} |\bar{x}_0 - \bar{x}_1|.$$

Proof. — The first part of the lemma follows immediately from the corresponding properties of Ψ and ϕ, using that the partial derivatives of θ of order one and two are bounded. The first statement in the second part of the lemma is an immediate consequence of the properties of C_w stated in the first part. One has also, from the first part

$$|C_w(w, \bar{x}) - 2(w - w^*)| \leqslant C\varepsilon_0 |w - w^*|, \tag{10.158}$$

$$(\mathbf{10.159}) \qquad |C(w,\bar{x}) + \delta(\bar{x}) - (w-w^*)^2| \leqslant C\varepsilon_0 |w-w^*|^2.$$

To prove the inequality in the second part of the lemma, the upper bound is just a consequence from the fact that, for each w, $C(w,\bar{x})$ is a Lipschitz function of $\bar{x}$. For the lower bound, we first recall that the partial derivatives θ_x, θ_y do not vanish and are bounded away from 0. Assume for instance that $\theta_x > C^{-1}$. Let w^* be the point where $C(w,\bar{x})$ attains its minimal value $-\delta(\bar{x})$. We have, from Lemma 18,

$$\phi(w^*,\bar{x}') \geqslant \phi(w^*,\bar{x}') + C^{-1}(\bar{x}'-\bar{x}),$$

hence

$$\begin{aligned} C(w^*,\bar{x}') &= (w^*)^2 - \theta(\Psi(w^*), \phi(w^*,\bar{x}')) \\ &\leqslant (w^*)^2 - \theta(\Psi(w^*), \phi(w^*,\bar{x})) - C^{-1}(\bar{x}'-\bar{x}) \\ &= -\delta(\bar{x}) - C^{-1}(\bar{x}'-\bar{x}), \end{aligned}$$

and the lower bound follows.

In the setting of the third part of the lemma, we have

$$(\mathbf{10.160}) \qquad C(w_0,\bar{x}_0) - C(w_1,\bar{x}_0) = C(w_1,\bar{x}_1) - C(w_1,\bar{x}_0),$$

where, by an argument just seen above,

$$(\mathbf{10.161}) \qquad |C(w_1,\bar{x}_1) - C(w_1,\bar{x}_0)| \geqslant C^{-1}|\bar{x}_0 - \bar{x}_1|.$$

On the other hand, we have

$$(\mathbf{10.162}) \qquad |C(w_0,\bar{x}_0) - C(w_1,\bar{x}_0)| \leqslant C|w_0-w_1| \max_w |C_w(w,\bar{x}_0)|,$$

where the maximum is taken for w between w_0 and w_1. But it follows from the first part of the lemma (second inequality) that this maximum is taken at w_0 (because we have assumed that $\delta(\bar{x}_0) > \delta(\bar{x}_1) > 0$), and then from (10.158), (10.159) that $|C_w(w_0,\bar{x}_0)| \leqslant C\delta(\bar{x}_0)^{1/2}$. Plugging this above completes the proof of the third part of the lemma. □

10.10.4. We now come back to the proof of the estimate (10.146). Let γ_0 be a horizontal closed segment in some rectangle R_a. Clearly, we may assume that $\mu([\gamma_0]) > 0$. By shortening γ_0 if necessary, we can assume that there is a stable curve through each endpoint of γ_0. Let $(P_0, Q_0, n_0) \in \mathcal{R}$ the element such that P_0 is the thinnest rectangle containing any stable curve in $[\gamma_0]$. There are at least two children of P_0 which contain a stable curve in $[\gamma_0]$.

We say that γ_0 has complexity 0 if at least one of the following two conditions are satisfied:

– At least one stable curve in $[\gamma_0]$ is contained in a simple child of P_0.

– There is a child P'_0 of P_0 such that the set of stable curves in $[\gamma_0]$ contained in P'_0 has μ-measure $\geqslant \frac{1}{3}\mu([\gamma_0])$.

If none of these conditions, we say that γ_0 has complexity > 0.

When γ_0 has complexity 0, it is easy to obtain (10.146) (and even better). Assume first that the first condition is satisfied. Then, as γ_0 intersects at least two children of P_0, we have

$$\text{(10.163)} \qquad \operatorname{diam} \gamma_0 \geqslant C^{-1}|P_0|.$$

On the other hand, we have

$$\text{(10.164)} \qquad \mu([\gamma_0]) \leqslant \mu(\{\omega \subset P_0\}) \leqslant C|P_0|^{d_s}$$

by Proposition 56, which proves (10.146).

Assume now that the second condition is satisfied. Then, as γ_0 intersects at least two children of P_0, we have now

$$\text{(10.165)} \qquad \operatorname{diam} \gamma_0 \geqslant C^{-1}|P'_0|.$$

On the other hand, we have

$$\text{(10.166)} \qquad \mu([\gamma_0]) \leqslant 3\mu(\{\omega \subset P'_0\}) \leqslant C|P'_0|^{d_s}$$

by Proposition 56, which proves again (10.146).

We have shown that $A(\gamma_0)$ is bounded for segments of complexity 0.

10.10.5. We now assume that the complexity of γ_0 is > 0. If some stable curve through an endpoint of γ_0 is contained in a child P'_0 of P_0, but there is a stable curve contained in P'_0 not in $[\gamma_0]$, we shorten γ_0 to remove from $[\gamma_0]$ all stable curves contained in P'_0. We do this for both endpoints. The shortened curve, that we denote by γ'_0, satisfies $\mu([\gamma'_0]) \geqslant \frac{1}{3}\mu([\gamma_0])$ (because the second condition above is not satisfied). The negation of the second condition also implies that at least two children of P_0 contain a curve in $[\gamma'_0]$. We can still assume there is a stable curve through each endpoint of γ'_0. Any child P'_0 of P_0 which contains a curve in $[\gamma'_0]$ is non-simple, and all stable curves contained in P'_0 belong to $[\gamma'_0]$.

In particular, P_0 has non-simple children, hence Q_0 is contained in Q_u. By (R4), $g_t^{n_0}(\gamma'_0)$ is a graph $\{y_u = \psi(x_u)\}$ satisfying (10.155). As there is a stable curve through each endpoint of γ'_0, the images of these endpoints by $g_t^{n_0}$ are contained in L_u; but then (10.155) implies that $g_t^{n_0}(\gamma'_0)$ is contained in L_u. Let $J_1 \subset J$ be the compact interval image by the projection on the second coordinate $\overline{x}$ of the curve $G \circ g_t^{n_0}(\gamma'_0)$.

We define $\delta_0 := \max_{J_1} \delta(\overline{x})$, $\delta'_0 := \min_{J_1} \delta(\overline{x})$. From Lemma 19, part 3, we have

$$\text{(10.167)} \qquad \delta_0^{1/2} \operatorname{diam} G \circ g_t^{n_0}(\gamma'_0) \geqslant C^{-1}|J_1|,$$

from which it follows that

(**10.168**) $$\delta_0^{1/2} \operatorname{diam} \gamma_0' \geqslant C^{-1} |P_0||J_1|.$$

We denote by $(P'_{0,i})_i$ the (non-simple) children of P_0 which contain a stable curve in $[\gamma_0]$. Each $P'_{0,i}$ is obtained from its parent P_0 by parabolic composition:

(**10.169**) $$(P'_{0,i}, Q'_{0,i}, n'_{0,i}) \in (P_0, Q_0, n_0) \,\square\, (P_{0,i}, Q_{0,i}, n_{0,i}).$$

Observe that it is possible that two $P'_{0,i}$ (but no more than two) correspond to the same $P_{0,i}$. The widths are related through

(**10.170**) $$C^{-1} \leqslant |P'_{0,i}||P_0|^{-1}|P_{0,i}|^{-1}\delta(Q_0, P_{0,i})^{\frac{1}{2}} \leqslant C.$$

Here, $\delta(Q_0, P_{0,i})$ is the quantity of Section 3.5.

For each i, we choose a stable curve $\{x_s = \varphi(y_s, \bar{x}_i)\}$ contained in $P_{0,i}$ and such that $\delta(\bar{x}_i)$ is not of the form $\delta_0 2^{-l}$. We have $\bar{x}_i \in J_1$ by construction of γ_0'. We also have (cf. Section 3.6.3)

(**10.171**) $$|\delta(Q_0, P_{0,i}) - \delta(\bar{x}_i)| \leqslant C(|Q| + |P_{0,i}|),$$

and therefore, in view of (R7) in Section 5.4

(**10.172**) $$\delta(\bar{x}_i) \gg (|Q| + |P_{0,i}|),$$

(**10.173**) $$C^{-1} \leqslant |P'_{0,i}||P_0|^{-1}|P_{0,i}|^{-1}\delta(\bar{x}_i)^{\frac{1}{2}} \leqslant C.$$

We now distinguish two cases

– $\delta_0' < \frac{1}{2}\delta_0$.
For any nonnegative integer l, let $J^*_{1,l}$ be the set of $\bar{x} \in J$ such that $\delta_0 2^{-l} \geqslant \delta(\bar{x}) \geqslant \delta_0 2^{-l-1}$. By Lemma 19, part 2, it is a compact interval satisfying

(**10.174**) $$|J^*_{1,l}| \leqslant C\delta_0 2^{-l}.$$

Let $\mathcal{L}$ be the set of nonnegative integers l such that $J^*_{1,l}$ contains some $\bar{x}_i$. It follows from (10.172) that $\mathcal{L}$ is finite. From (10.172), it also follows that it is possible to find, for $l \in \mathcal{L}$, an interval $J_{1,l}$ whose endpoints are at a distance $\ll \delta_0 2^{-l}$ from those of $J^*_{1,l}$ and which has the following property: if $\bar{x}_i \in J^*_{1,l}$, for any stable curve $\varphi(y_s, \bar{x})$ contained in $P_{0,i}$, one has $\bar{x} \in J_{1,l}$. One has still

(**10.175**) $$|J_{1,l}| \leqslant C\delta_0 2^{-l}.$$

– $\delta_0' \geqslant \frac{1}{2}\delta_0$.
In this case, we set $\mathcal{L} = \{0\}$, $J_{1,0} = J_1$.

In both cases, for $l \in \mathcal{L}$, let $\gamma_{1,l} \subset \mathcal{J}$ be the horizontal segment $J_{1,l} \times \{y_s^*\}$.

10.10.6. In the first case, we write

$$\mu([\gamma_0]) \leqslant 3\mu([\gamma_0']) \leqslant C\sum_i |P'_{0,i}|^{d_s} \quad \text{(from Proposition 56)}$$

$$\leqslant C|P_0|^{d_s}\sum_i |P_{0,i}|^{d_s}\delta(\overline{x}_i)^{-\frac{1}{2}d_s} \quad \text{(from (10.173))}$$

$$\leqslant C|P_0|^{d_s}\delta_0^{-\frac{1}{2}d_s}\sum_{\mathcal{L}} 2^{\frac{lds}{2}}\sum^{(l)} |P_{0,i}|^{d_s}$$

$$\leqslant C|P_0|^{d_s}\delta_0^{-\frac{1}{2}d_s}\sum_{\mathcal{L}} 2^{\frac{lds}{2}}\sum^{(l)} \mu(\{\omega\subset P_{0,i}\})$$

(from Proposition 56)

$$\leqslant C|P_0|^{d_s}\delta_0^{-\frac{1}{2}d_s}\sum_{\mathcal{L}} 2^{\frac{lds}{2}}\mu([\gamma_{1,l}]). \tag{10.176}$$

We have written $\sum^{(l)}$ for the partial sum over those $P_{0,i}$ satisfying $\overline{x}_i\in J^*_{1,l}$.

When $\delta_0'\geqslant\frac{1}{2}\delta_0$, a similar but simpler argument gives

$$\mu([\gamma_0]) \leqslant C|P_0|^{d_s}\delta_0^{-\frac{1}{2}d_s}\mu([\gamma_{1,0}]). \tag{10.177}$$

By definition of $A(\gamma)$, we have, for all l

$$\mu([\gamma_{1,l}]) = A(\gamma_{1,l})(\operatorname{diam}\gamma_{1,l})^{d_s}. \tag{10.178}$$

When $\delta_0'\geqslant\frac{1}{2}\delta_0$, we now use (10.168), (10.177) to conclude that

$$A(\gamma_0) \leqslant CA(\gamma_{1,0}). \tag{10.179}$$

When $\delta_0'<\frac{1}{2}\delta_0$, we use (10.176), (10.175), (10.168) to obtain

$$\mu([\gamma_0]) \leqslant C\max_{\mathcal{L}} A(\gamma_{1,l})|P_0|^{d_s}\delta_0^{\frac{1}{2}d_s}\sum_{\mathcal{L}} 2^{\frac{-lds}{2}}$$

$$\leqslant C\max_{\mathcal{L}} A(\gamma_{1,l})|P_0|^{d_s}\delta_0^{\frac{1}{2}d_s}$$

$$\leqslant C\max_{\mathcal{L}} A(\gamma_{1,l})(\delta_0|J_1|^{-1}\operatorname{diam}\gamma_0)^{d_s}. \tag{10.180}$$

But in this case, we have $|J_1|\geqslant C^{-1}\delta_0$ from Lemma 19, part 2, and we conclude that

$$A(\gamma_0) \leqslant C\max_{\mathcal{L}} A(\gamma_{1,l}). \tag{10.181}$$

To obtain (10.146), it is thus sufficient to define, for all horizontal segments γ_0 with $\mu([\gamma_0]) > 0$, a complexity index $c(\gamma_0) \in \mathbf{N}$ which vanishes as explained in Section 10.10.4 and satisfies otherwise

$$c(\gamma_0) \leqslant C \log\log |P_0|^{-1}, \tag{10.182}$$

$$c(\gamma_0) = 1 + \max_{\mathcal{L}} c(\gamma_{1,l}). \tag{10.183}$$

10.10.7. We want to use (10.183) to give an inductive definition of $c(\gamma_0)$. This will work if the $\gamma_{1,l}$ are in some sense "simpler" than γ_0. If all $\gamma_{1,l}$ have complexity 0, we just set $c(\gamma_0) = 1$. Assume therefore that some $\gamma_{1,l}$, that we just denote by γ_1, has complexity > 0 (according to Section 10.10.4). Observe that $J^*_{1,l}$ must contain at least two $\bar{x}_i$: otherwise, the unique $P_{0,i}$ such that $\bar{x}_i \in J^*_{1,l}$ would be the thinnest rectangle containing any stable curve in $[\gamma_1]$, and γ_1 would have complexity 0 by the first condition of Section 10.10.4.

Therefore, there exists an element $(P_1, Q_1, n_1) \in \mathcal{R}$ with the following properties:

- each $P_{0,i}$ with $\delta(\bar{x}_i) \in J^*_{1,l}$ is contained in some non-simple child of P_1;
- at least two non-simple children of P_1 contain some such $P_{0,i}$.

Lemma **20.** — *One has*

$$|P_1|^{1-\eta} \geqslant C^{-1}\delta_0 2^{-l}.$$

Moreover P_1 *is* I-*critical for any parameter interval* I *containing* t.

Proof. — Let i such that $P_{0,i} \subset P_1$. There exists a parameter interval I containing t such that Q_0 and $P_{0,i}$ are I-transverse. On the other hand, as $P'_{0,i}$ is a child of P_0, Q_0 and P_1 are I-critically related for every parameter interval I containing t. By Proposition 21 in Section 8.1, we must have $|P_1| > \frac{1}{3}|Q_0|$. If we had $|P_1|^{1-\eta} \ll \delta_0 2^{-l}$, we would also have $|Q_0|^{1-\eta} \ll \delta_0 2^{-l}$ and (T1), (T2), (T3) for $Q_0 \,\overline{\pitchfork}_I\, P_1$ would hold for a sufficiently small parameter interval I containing t. This contradiction proves the first statement. The second follow from Proposition 24 in Section 8.2. □

10.10.8. We assume in this subsection that P_0 also is I-critical for any I containing t.

Let I^* be the largest parameter interval for which we have

$$|P_0| > |I^*|^{\beta}. \tag{10.184}$$

Observe that P_0 is I^*-defined by Corollary 6 in Section 6.6.3. Then, (P_0, Q_0, n_0) cannot be I^*-bicritical as I^* is β-regular, hence Q_0 is I^*-transverse. This implies that there exists $P_{0,i}$ with $\delta(\bar{x}_i) \in J^*_{1,l}$ and P^* containing the stable curve $\{x_s = \varphi(y_s, \bar{x}_i)\}$ such that Q_0 and P^* are I^*-transverse. From Section 3.6 and Lemma 3 in Section 6.6.3, we have

$$\delta_0 2^{-l} \geqslant \delta(\bar{x}_i) \geqslant \delta(Q_0, P^*) \geqslant |I^*|. \tag{10.185}$$

On the other hand, it follows from the definition of I^* that

$$|I^*|^\beta \geqslant |P_0|^{1+\tau}. \tag{10.186}$$

Inequalities (10.185), (10.186), and Lemma 20 give

$$|P_0| < |P_1|^{\frac{1}{2}(1+\beta)}. \tag{10.187}$$

This estimate means that every P_1 that occur (for the various $\gamma_1 = \gamma_{1,l}$ such that the complexity of $\gamma_{1,l}$ is > 0) is indeed simpler than P_0. Moreover, by Lemma 20, the hypothesis of Section 10.10.8 is satisfied by any such P_1. We can therefore under the assumption of Section 10.10.8 use (10.183) to define inductively $c(\gamma_0)$. The inequality (10.182) follows then from (10.187).

10.10.9. Finally we deal with the general case. By Lemma 20, the index $c(\gamma_{1,l})$ is already defined for every $l \in \mathcal{L}$. We define $c(\gamma_0)$ by (10.183) and check (10.182). For every $l \in \mathcal{L}$, we have either $c(\gamma_{1,l}) = 0$ or $c(\gamma_{1,l}) > 0$ and

$$c(\gamma_{1,l}) \leqslant C \log\log |P_1|^{-1}, \tag{10.188}$$

where P_1 is associated to $\gamma_{1,l}$ as in Section 10.10.7.

From Proposition 12 in Section 6.6.2, we have

$$\log\log |P_0|^{-1} \geqslant C^{-1} \log n_0. \tag{10.189}$$

On the other hand, we have, from Lemma 20 and (10.172)

$$C^{-1} \log |P_1|^{-1} \leqslant \log(\delta_0 2^{-l}) \leqslant \log |Q_0|^{-1} \leqslant C n_0. \tag{10.190}$$

From this, we obtain

$$\log\log |P_1|^{-1} \leqslant \log\log |P_0|^{-1} + C, \tag{10.191}$$

and (10.182) follows.

The proof of the theorem is now complete.

10.11. *Invariant measures.* — From the Gibbs measure μ, which is not invariant but has the Jacobian property, we define a measure ν on $\mathcal{R}_+^\infty$ by $d\nu = h^* d\mu$. Recall that $\mu(\mathcal{R}_+^\infty - \mathcal{D}_+^\infty) = 0$ and that h^* is normalized by $\mu(h^*) = 1$. This means that ν is a probability measure on $\mathcal{R}_+^\infty$ and that $\nu(\mathcal{D}_+^\infty) = 1$.

Proposition **58.** — *The probability measure $d\nu = h^* d\mu$ is T^+-invariant, ergodic. It satisfies, for all $(P, Q, n) \in \mathcal{R}$:*

$$C^{-1} |P|^{d_s} \leqslant \nu(\{\omega \subset P\}) \leqslant C |P|^{d_s}.$$

Proof. — As h^* is bounded and bounded away from 0 (cf. (10.116)), the estimate for $\nu(\{\omega \subset P\})$ follows from the same estimate for $\mu(\{\omega \subset P\})$ in Proposition 56.

Let us prove that ν is 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

$$|h_0 h_1|_E \leqslant \|h_0\|_\infty |h_1|_E + |h_0|_E \|h_1\|_\infty. \tag{10.192}$$

In particular, for any $h \in E$, hh^* also belongs to E. Let $h \in E$. We write

$$\begin{aligned} \int h(T^+\omega)d\nu(\omega) &= \int h(T^+\omega)h^*(\omega)d\mu(\omega) \\ &= \sum_{\mathcal{P}} \int h(T^+\omega)h^*(\omega)\chi_P^*(\omega)d\mu(\omega), \end{aligned} \tag{10.193}$$

where χ_P^* is the characteristic function of $\mathcal{R}_+^\infty(P)$. The Jacobian property (10.123) (Section 10.9) gives

$$\begin{aligned} &\int h(T^+\omega)h^*(\omega)\chi_P^*(\omega)d\mu(\omega) \\ &= \int_{\mathcal{R}_+^\infty(a)} h(\omega)h^*(T_P^{-1}(\omega))\exp(-d_s b(T_P^{-1}(\omega)))d\mu(\omega) \end{aligned} \tag{10.194}$$

where $Q \subset R_a$ and T_P is the restriction of T^+ to $\mathcal{R}_+^\infty(P)$. Summing over P and using that h^* is L_{d_s}-invariant gives

$$\int h(T^+\omega)d\nu(\omega) = \int h(\omega)d\nu(\omega). \tag{10.195}$$

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 prove 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 $a \in \mathcal{a}$ such that

$$\nu(A \cap \mathcal{R}_+^\infty(a)) \geqslant (1-\varepsilon)\nu(\mathcal{R}_+^\infty(a)). \tag{10.196}$$

As $\varepsilon > 0$ is arbitrary, this easily implies $\nu(A) = 1$.

As $\nu(A) > 0$, we can find (P, Q, n) such that

$$\nu(\{\omega \subset P\} \cap A^c) \leqslant \varepsilon' \nu(\{\omega \subset P\}), \tag{10.197}$$

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

$$\{\omega \subset P\} = \bigcup_{0 \leqslant j \leqslant r} \bigcup_{P_j} (T^+)^{-j}(\mathcal{R}_+^\infty(P_j)) \bmod 0 \tag{10.198}$$

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 (10.197), there exists $0 \leqslant j \leqslant r$ and P_j such that

$$(\mathbf{10.199}) \qquad \nu(A^c \cap (T^+)^{-j}(\mathcal{R}_+^\infty(P_j))) \leqslant \varepsilon' \nu((T^+)^{-j}(\mathcal{R}_+^\infty(P_j))).$$

We apply the Jacobian property, taking (10.118) into account to get (10.196) with $\varepsilon = C\varepsilon'$. We have proved that ν is ergodic, and the proof of the proposition is complete. □

We will now lift ν to obtain a $\widetilde{T}^+$-invariant probability measure on $\widetilde{\mathcal{R}}_+^\infty$.

Proposition **59.** — *There exists a unique probability measure* $\widetilde{\nu}$ *on* $\widetilde{\mathcal{R}}_+^\infty$ *which is* $\widetilde{T}^+$*-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 $\widetilde{\mathcal{R}}_+^\infty$ which project onto ν. This is a compact set for the weak topology, invariant under $\widetilde{T}^+$ because ν is T^+-invariant. One obtains a $\widetilde{T}^+$-invariant measure in $\mathcal{M}(\nu)$ by taking any $\widetilde{\nu}_0 \in \mathcal{M}(\nu)$ and choosing a weak limit of a subsequence of

$$(\mathbf{10.200}) \qquad \frac{1}{n}\sum_0^{n-1}[(\widetilde{T}^+)^j]^*(\widetilde{\nu}_0).$$

Uniqueness. — The set of fixed points for the action of $\widetilde{T}^+$ 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 $\widetilde{\nu}_0, \widetilde{\nu}_1$. As ν is ergodic, $\widetilde{\nu}_0$ and $\widetilde{\nu}_1$ are also ergodic. Still by the ergodicity of ν, some stable curve ω must meet the basins of both $\widetilde{\nu}_0$ and $\widetilde{\nu}_1$. But stable curves are contracted exponentially fast under positive iteration by T^+; we should thus have $\widetilde{\nu}_0 = \widetilde{\nu}_1$, a contradiction.

We have already said that $\widetilde{\nu}$ is ergodic. □

Finally, we want to "spread" the $\widetilde{T}^+$-invariant measure $\widetilde{\nu}$ in order to obtain a g-invariant measure σ. Let $\Lambda = \Lambda_g$ as in the Introduction (cf. Section 1.2).

We first observe that the support of $\widetilde{\nu}$ is contained into $\Lambda \cap \widetilde{\mathcal{R}}_+^\infty$: if $N \subset \widetilde{\mathcal{R}}_+^\infty$ is compact and disjoint from Λ, then N is disjoint from the image of $(\widetilde{T}^+)^j$ if j is large enough, hence $\widetilde{\nu}(N) = 0$.

Let now h be a continuous, and thus bounded, function on Λ. For $x \in \Lambda \cap \mathcal{D}_+^\infty$, we write

$$(\mathbf{10.201}) \qquad \widetilde{T}^+(x) = g^{N(x)}(x),$$

where $N(x) = n$ if $x \in \widetilde{\mathcal{R}}^{\infty}_{+}(P)$ with $(P, Q, n) \in \mathcal{P}$. We define:

$$\textbf{(10.202)} \qquad Sh(x) = \sum_{0 \leqslant j < N(x)} h(g^j(x)).$$

The function Sh is defined $\widetilde{\nu}$-almost everywhere. It satisfies:

$$\textbf{(10.203)} \qquad |Sh(x)| \leqslant \|h\|_{\infty} N(x).$$

By Proposition 58 and Corollary 16 in Section 10.3, the function N is $\widetilde{\nu}$-integrable. We have therefore defined an operator

$$\textbf{(10.204)} \qquad S : C(\Lambda) \mapsto L^1(\widetilde{\nu}),$$

where $C(\Lambda)$ stands for the space of continuous functions on $\Lambda = \Lambda_g$.

We define a finite measure σ on Λ by

$$\textbf{(10.205)} \qquad \int h d\sigma = \int Sh d\widetilde{\nu},$$

for $h \in C(\Lambda)$. From the definition of Sh, we have

$$\textbf{(10.206)} \qquad S(h \circ g) = Sh + h \circ \widetilde{T}^+ - h.$$

Thus, the $\widetilde{T}^+$-invariance of $\widetilde{\nu}$ implies that σ is g-invariant. It is ergodic. The Lyapunov exponents of $\widetilde{T}^+$ for $\widetilde{\nu}$ are non-zero because $\widetilde{T}^+$ is uniformly hyperbolic. To get the Lyapunov exponents of g for σ we have only to change time, which is possible since N is $\widetilde{\nu}$-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 $\mathcal{R}^{\infty}_{-}$, as defined in Section 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 Section 10.4. Let $(\mathrm{P}_k^*, \mathrm{Q}_k^*, n_k^*)_{k\geqslant 0}$ be the canonical sequence associated to ω^* (cf. definition also in Section 10.4). We have

$$\omega^* = \bigcap_{k\geqslant 0} \mathrm{Q}_k^*, \tag{11.1}$$

where Q_0^* is a rectangle R_a and Q_{k+1}^* is a child of Q_k^* for each $k \geqslant 0$. We want to analyze the intersection $\omega^* \cap \Lambda$. In Section 10, we have analyzed the set $\widetilde{\mathcal{R}}_+^\infty$ and we know, in particular, that $\omega^* \cap \Lambda$ contains the subset $\omega^* \cap \widetilde{\mathcal{R}}_+^\infty$; 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 **5.** — *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 $\mathrm{C}r(\omega^*)$,*
– *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 $\omega^*{}_i$ and an integer n_i such that*

$$g^{n_i}(\omega^*(i)) = \omega_i^*, \tag{11.2}$$

$$g^{n_i}(\Lambda_i(\omega^*)) = \omega_i^* \cap \widetilde{\mathcal{R}}_+^\infty. \tag{11.3}$$

In particular, there is a special index $i = 0$ for which $n_0 = 0$, $\omega^*(0) = \omega_0^* = \omega^*$, $\Lambda_0(\omega^*) = \omega^* \cap \widetilde{\mathcal{R}}_+^\infty$.

(ii) *For every point $c \in \mathrm{C}r(\omega^*)$, 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 $g^{\mathrm{N}_0}(\omega^-(c) \cap \mathrm{L}_u)$.*

(iii) *The Hausdorff dimension of $\mathcal{E}(\omega^*)$ is not greater than*

$$\left(\frac{1}{d_s^0} + \frac{1}{2(d_s^0 + d_u^0 - 1)}\right)^{-1} + o(1) \tag{11.4}$$

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 $(\omega^*(i_n(x)))_{n\geqslant 0}$.*

Remark **18.**

1. The structure of $\omega^* \cap \Lambda$ 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, $\mathrm{C}r(\omega^*)$ 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 $\mathrm{C}r(\omega^*)$ can be empty.

11.2. *Parabolic cores.* — Let $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}$, $\mathcal{R}$ as in Section 10.1.

Definition **16.** — *The parabolic core of* P, *denoted by* $c(\mathrm{P})$, *is the set of points of* $\mathrm{W}^s(\Lambda, \widehat{\mathrm{R}})$ *which belong to* P *but not to any child of* P. *The parabolic core of* Q, *denoted by* $c(\mathrm{Q})$, *is the set of points of* $\mathrm{W}^u(\Lambda, \widehat{\mathrm{R}})$ *which belong to* Q *but not to any child of* Q.

We have partitions

$$\mathrm{R} \cap \mathrm{W}^s(\Lambda, \widehat{\mathrm{R}}) = \bigsqcup_{\mathcal{R}} c(\mathrm{P}) \sqcup \widetilde{\mathcal{R}}_+^{\infty}, \tag{11.5}$$

$$\mathrm{R} \cap \mathrm{W}^u(\Lambda, \widehat{\mathrm{R}}) = \bigsqcup_{\mathcal{R}} c(\mathrm{Q}) \sqcup \widetilde{\mathcal{R}}_-^{\infty}. \tag{11.6}$$

If R_a is the rectangle which contains ω^*, we also have

$$\omega^* \cap \Lambda = \bigsqcup_{\mathrm{P} \subset \mathrm{R}_a} (\omega^* \cap c(\mathrm{P})) \sqcup (\omega^* \cap \widetilde{\mathcal{R}}_+^{\infty}). \tag{11.7}$$

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, $c(\mathrm{P})$ is empty if Q is I-transverse. Thus, the union in (11.5), (11.7) can be restricted to those $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}$ such that Q is I-critical for all I.

We will denote by $\mathrm{C}(\omega^*)$ the set of elements $(\mathrm{P}, \mathrm{Q}, n) \in \mathcal{R}$ such that $c(\mathrm{P}) \cap \omega^*$ is not empty. For any $(\mathrm{P}, \mathrm{Q}, n) \in \mathrm{C}(\omega^*)$, Q is I-critical for all I.

11.3. *Decomposition of* $c(\mathrm{P}) \cap \omega^*$. — Let $(\mathrm{P}, \mathrm{Q}, n) \in \mathrm{C}(\omega^*)$. For $k \geqslant 0$, set

$$(\mathrm{P}_k, \mathrm{Q}_k, n_k) = (\mathrm{P}_k^*, \mathrm{Q}_k^*, n_k^*) * (\mathrm{P}, \mathrm{Q}, n), \tag{11.8}$$

$$\omega_{\mathrm{P}}^* = \bigcap_{k \geqslant 0} \mathrm{Q}_k. \tag{11.9}$$

The unstable curve ω_{P}^* is contained in Q and we have

$$g^n(\omega^* \cap c(\mathrm{P})) \subset \omega_{\mathrm{P}}^* \cap \mathrm{L}_u. \tag{11.10}$$

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 $(P, Q, n) \boxdot (P', Q', n')$ is well defined and produces two children of P.

Let P' be a transverse vertex of level $k > 0$. For all $m \geqslant k$, the parabolic composition

$$(P_m, Q_m, n_m) \boxdot (P', Q', n')$$

is well-defined and produces two elements $(P^\pm_m, Q^\pm_m, n^\pm_m)$. The formulas

$$\omega^*_{P,P',+} := \bigcap Q^+_m, \qquad \omega^*_{P,P',-} := \bigcap Q^-_m, \tag{11.11}$$

define unstable curves $\omega^*_{P,P',\pm}$ contained in Q'. We also define pieces $\omega^*(P, P', \pm)$ of ω^* through

$$g^{n_{P,P'}}(\omega^*(P, P', \pm)) = \omega^*_{P,P',\pm}, \tag{11.12}$$

$$n_{P,P'} := n + n' + N_0. \tag{11.13}$$

Lemma **21.** — *Let x be a point in $\omega^* \cap c(P)$, $y = g^{n+N_0}(x)$. Either y belongs 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 **60.** — *There is at most one point $x \in \omega^* \cap c(\mathrm{P})$ such that $y = g^{n+\mathrm{N}_0}(x)$ 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 $g^{\mathrm{N}_0}(\mathrm{L}_u \cap \omega^*_{\mathrm{P}})$ at y as a quadratic tangency point.*

Proof. — Let x be a point in $\omega^* \cap c(\mathrm{P})$ such that $y = g^{n+\mathrm{N}_0}(x)$ belongs to a decreasing sequence $(\mathrm{P}'_\ell)_{\ell \geqslant 0}$ of critical vertices. Denote by ω_+ the stable curve which is the intersection of these critical vertices. For all parameter intervals I, all $k \geqslant 0$, $\ell \geqslant 0$, Q_k and P'_ℓ are I-critically related. This implies that

$$\lim_{\substack{k \to +\infty \\ \ell \to +\infty}} \delta(\mathrm{Q}_k, \mathrm{P}'_\ell) = 0. \tag{11.14}$$

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 $g^{\mathrm{N}_0}(\gamma_k)$ goes to zero as k, ℓ go to $+\infty$. Passing to the limit, we see that ω_+ has a tangency with $g^{\mathrm{N}_0}(\omega^*_{\mathrm{P}} \cap \mathrm{L}_u)$. This tangency is quadratic in the following sense (cf. also the remark after the end of the proof): First, $g^{\mathrm{N}_0}(\omega^*_{\mathrm{P}} \cap \mathrm{L}_u)$ is contained, with the exception of the tangency point, in one of the components of $\mathrm{P}_s - \omega_+$; moreover, the angle between the tangent lines to $\omega_+(x)$, $g^{\mathrm{N}_0}(\mathrm{L}_u \cap \omega^*_{\mathrm{P}})$ 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 Section 3.5.

As ω_+ and $g^{\mathrm{N}_0}(\mathrm{L}_u \cap \omega^*_{\mathrm{P}})$ 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 $g^{\mathrm{N}_0}(\mathrm{L}_u \cap \omega^*_{\mathrm{P}}) \cap \omega_+$ or $g^{\mathrm{N}_0}(\mathrm{L}_u \cap \omega^*_{\mathrm{P}}) \cap \omega'_+$ is empty. But, then, we have $y' := g^{n+\mathrm{N}_0}(x') = y$ and $x' = x$. □

Remark **19.** — 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 $\mathrm{C}^{1+\mathrm{Lip}}$ regularity class, where the notion of "quadratic" tangency, as explained in the proof of Proposition 60, still makes sense.

It is easy to see exactly when a point $x \in \omega^* \cap c(\mathrm{P})$ with the property specified in Proposition 60 does exist: a necessary and sufficient condition is that the tree $\mathcal{A}(\omega^*, \mathrm{P})$ is infinite. In this case, the point x will be a point of the set $\mathrm{C}r(\omega^*)$ in the statement of Theorem 5 and the point $y = g^{n+\mathrm{N}_0}(x)$ is said to be *critical.*

Summarizing what we have established so far, two cases may happen:

(1) The tree $\mathcal{A}(\omega^*, \mathrm{P})$ is finite. Then, the intersection $\omega^* \cap c(\mathrm{P})$ is the finite disjoint union of the sets

$$\omega^*(\mathrm{P}, \mathrm{P}', \pm) \cap \Lambda \tag{11.15}$$

where P′ runs through the vertices of the tree which are transverse of level > 0. The image under $g^{n_{\mathrm{P},\mathrm{P}'}}$ of the set (11.15) is the intersection $\omega^*_{\mathrm{P},\mathrm{P}',\pm} \cap \Lambda$.

(2) The tree $\mathcal{A}(\omega^*, \mathrm{P})$ is infinite. Then, the intersection $\omega^* \cap c(\mathrm{P})$ is the countable disjoint union of the sets $\omega^*(\mathrm{P}, \mathrm{P}', \pm) \cap \Lambda$ as above and a single point $x \in \mathrm{C}r(\omega^*)$. The point $x = x_{\mathrm{P}}$ is the limit of the pieces $\omega^*(\mathrm{P}, \mathrm{P}', \pm)$ (whose diameters goes to 0 as $|\mathrm{P}'|$ goes to 0).

11.4. *The structure of* $\omega^* \cap \Lambda$. — We are now ready to prove all the statements in Theorem 5, mentioned above in Section 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 5, is obtained by iterating the partition (11.7) and the decomposition of $\omega^* \cap c(\mathrm{P})$ described in Section 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 $(\mathrm{P}, \mathrm{Q}, n) \in \mathrm{C}(\omega^*)$ such that $\mathcal{A}(\mathrm{P}, \omega^*)$ is infinite, a point x_{P} such that $y_{\mathrm{P}} = g^{n+\mathrm{N}_0}(x_{\mathrm{P}})$ is critical; such points x_{P} are said of type II;
- for each $(\mathrm{P}, \mathrm{Q}, n) \in \mathrm{C}(\omega^*)$, each vertex $(\mathrm{P}', \mathrm{Q}', n')$ of $\mathcal{A}(\omega^*, \mathrm{P})$ which is transverse of level bigger than 0, each $\varepsilon \in \{+, -\}$, the intersection $\omega^*(\mathrm{P}, \mathrm{P}', \varepsilon) \cap \Lambda$; the image of this set under $g^{n_{\mathrm{P},\mathrm{P}'}}$ is the intersection $\omega^*_{\mathrm{P},\mathrm{P}',\varepsilon} \cap \Lambda$ of another unstable curve with Λ.

The intersection $\omega^*_{\mathrm{P},\mathrm{P}',\varepsilon} \cap \Lambda$ 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 5. If it is of type II, it belongs to $\mathrm{C}r(\omega^*)$. Assume now that it is of type III. Then, it belongs to some $\omega^*(\mathrm{P}, \mathrm{P}', \varepsilon) \cap \Lambda$ as above. Define

$$z_1 = g^{n_{\mathrm{P},\mathrm{P}'}}(z_0), \tag{11.16}$$

which belongs to $\omega^*_{\mathrm{P},\mathrm{P}',\varepsilon} \cap \Lambda =: \omega^*_1$. 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 $\omega^*_1(\mathrm{P}_1, \mathrm{P}'_1, \varepsilon_1)$; we define

$$z_2 = g^{n_{\mathrm{P}_1,\mathrm{P}'_1}}(z_1), \tag{11.17}$$

which belongs to $\omega^*_2 \cap \Lambda$, with

$$\omega^*_2 := g^{n_{\mathrm{P}_1,\mathrm{P}'_1}}(\omega^*_1(\mathrm{P}_1, \mathrm{P}'_1, \varepsilon)). \tag{11.18}$$

Iterating this process lead to one of three possible outcomes:

(1) the z_k's are defined and of type III for all $k \geqslant 0$; the corresponding initial points z_0 form the set $\mathcal{E}(\omega^*)$.
(2) the z_k's are defined for $0 \leqslant k \leqslant \ell$ and z_ℓ is of type I, i.e. it belongs to $\widetilde{\mathcal{R}}_+^\infty$; let $(P_k, P'_k, \varepsilon_k)$ for $0 \leqslant k < \ell$ 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 5.
(3) the z_k's are defined for $0 \leqslant k \leqslant \ell$ 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 5. The properties (i), (ii), (iv) follow immediately from the definitions.

11.5. *Hausdorff dimension of the exceptional set* $\mathcal{E}(\omega^*)$.

11.5.1. 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

$$(\mathbf{11.19}) \qquad \mathcal{E}(\omega^*) = \bigsqcup_{(P,P',\varepsilon)} g^{-n_{P,P'}}(\mathcal{E}({\omega^*}_{P,P',\varepsilon})),$$

where $\varepsilon \in \{+,-\}$, P runs through $C(\omega^*)$ and P′ through vertices of $\mathcal{A}(\omega^*, P)$ which are transverse of level > 0.

11.5.2. The estimate on the Hausdorff dimension will follow from a standard result that we formulate in a general setting.

Let $\alpha, d \in (0,1)$. Let Ω be a set. For each $\omega \in \Omega$, we are given

– a subset $E(\omega) \subset [0,1]$ which is the union of at most countably many disjoint compact subintervals of $[0,1]$;
– a map $F_\omega = (g_\omega, f_\omega)$ from $E(\omega)$ into $[0,1] \times \Omega$.

These data satisfy

– Each map f_ω is constant on each component of $E(\omega)$.
– The restriction of each map g_ω to each component of $E(\omega)$ is an uniformly expansive $C^{1+\alpha}$ diffeomorphism onto $[0,1]$ with uniformly bounded distortion.

The second condition means that there exists $0 < \lambda < 1$, $C_0 > 0$, independent of ω, such that, writing h_J for the inverse of the restriction of g_ω to a component J of $E(\omega)$, we have, for $x, y \in [0,1]$

$$|Dh_J(x)| \leqslant \lambda,$$
$$|\log|Dh_J(x)| - \log|Dh_J(y)|| \leqslant C_0|x-y|^\alpha.$$

Define $E := \{(x,\omega) \in [0,1] \times \Omega,\, x \in E(\omega)\}$ and the map $F : E \to [0,1] \times \Omega$ by $F(x,\omega) = F_\omega(x)$. Let $\mathcal{E} = \bigcap_{n\geqslant 0} F^{-n}(E)$. For $\omega \in \Omega$, let $\mathcal{E}(\omega) := \{x \in [0,1], (x,\omega) \in \mathcal{E}\}$.

Proposition **61.** — *Assume that one has, for every* $\omega \in \Omega$

$$\sum_{\mathrm{J}} |\mathrm{J}|^d \leqslant \exp(-\mathrm{C}_0 d(1-\lambda^\alpha)^{-1}),$$

where the sum runs over the components J *of* $\mathrm{E}(\omega)$. *Then the Hausdorff dimension of each set* $\mathcal{E}(\omega)$ *is at most* d.

Proof. — For $n > 0$, $\omega \in \Omega$, let $\mathrm{E}_n := \bigcap_{0 \leqslant m < n} \mathrm{F}^{-m}(\mathrm{E})$ be the domain of F^n, and $\mathrm{E}_n(\omega) := \{x \in [0,1], (x,\omega) \in \mathrm{E}_n\}$ be the fiber of E_n. Each $\mathrm{E}_n(\omega)$ is the union of at most countably many disjoint compact intervals.

Let $n > 0$, $\omega \in \Omega$, J be a component of $\mathrm{E}_n(\omega)$. There exists ω' and a $\mathrm{C}^{1+\alpha}$ diffeomorphism h_J from $[0,1]$ onto J such that $\mathrm{F}^n(x,\omega) = (h_\mathrm{J}^{-1}(x), \omega')$ for $x \in \mathrm{J}$. Moreover, we have, for $x, y \in [0,1]$

$$|\mathrm{D}h_\mathrm{J}(x)| \leqslant \lambda^n,$$
$$|\log|\mathrm{D}h_\mathrm{J}(x)| - \log|\mathrm{D}h_\mathrm{J}(y)|| \leqslant \mathrm{C}_0(1-\lambda^\alpha)^{-1}|x-y|^\alpha.$$

For $n > 0$, $\omega \in \Omega$, define

$$\mathrm{S}_n(\omega) = \sum_{\mathrm{J}} |\mathrm{J}|^d,$$

where the sum runs over the components J of $\mathrm{E}_n(\omega)$. Here, each interval has length $\leqslant \lambda^n$. Therefore the components of $\mathrm{E}_n(\omega)$ form a covering of $\mathcal{E}(\omega)$ by intervals of small diameter. We have by hypothesis $\mathrm{S}_1(\omega) \leqslant \exp(-\mathrm{C}_0 d(1-\lambda^\alpha)^{-1})$. We will show that $\mathrm{S}_{n+1}(\omega) \leqslant \mathrm{S}_n(\omega)$, which implies the conclusion of the proposition.

In the sum for $\mathrm{S}_{n+1}(\omega)$, we first sum over components J^* of $\mathrm{E}_{n+1}(\omega)$ which are contained in a fixed component J of $\mathrm{E}_n(\omega)$. With ω' as above, such J^* are exactly the images by h_J of the components J' of $\mathrm{E}(\omega')$. The lengths are related through the mean-value theorem by

$$|\mathrm{J}| = |\mathrm{D}h_\mathrm{J}(x)|, \qquad |\mathrm{J}^*| = |\mathrm{D}h_\mathrm{J}(y)||\mathrm{J}'|,$$

for some $x \in [0,1]$, $y \in \mathrm{J}'$. We therefore have

$$|\mathrm{J}^*| \leqslant \exp(\mathrm{C}_0(1-\lambda^\alpha)^{-1})|\mathrm{J}||\mathrm{J}'|.$$

We finally obtain

$$\begin{aligned}\mathrm{S}_{n+1}(\omega) &= \sum_{\mathrm{J}}\sum_{\mathrm{J}'} |\mathrm{J}^*|^d \\ &\leqslant \exp(\mathrm{C}_0 d(1-\lambda^\alpha)^{-1}) \sum_{\mathrm{J}}\Big(|\mathrm{J}|^d \sum_{\mathrm{J}'} |\mathrm{J}'|^d\Big)\end{aligned}$$

$$= \exp(C_0 d(1-\lambda^\alpha)^{-1}) \sum_J (|J|^d S_1(\omega'))$$

$$\leqslant \sum_J |J|^d = S_n(\omega).$$

□

11.5.3. We now come back to the setting of the theorem.

Lemma **22.** — *The maps*

$$g^{n_{P,P'}} : \omega^*(P, P', \varepsilon) \to \omega^*{}_{P,P',\varepsilon}$$

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

$$g^{n^*_k} : P^*_k \to Q^*_k, \qquad g^{n'_k} : P'_k \to Q'_k, \tag{11.20}$$

have bounded distortion, hence the one-dimensional map

$$g^{n^*_k} \circ (g^{n'_k})^{-1} : \gamma'_k \to \gamma_k \tag{11.21}$$

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. □

Lemma **23.** — *Let*

$$\delta(\omega^*_P, P') = \lim_{k\to+\infty} \delta(Q_k, P').$$

We have

$$C^{-1} \leqslant \frac{\operatorname{diam} \omega^*(P, P', \varepsilon)}{|P||P'|(\delta(\omega^*_P, P'))^{-\frac{1}{2}}} \leqslant C.$$

Proof. — As in the proof of Lemma 22, we write

$$g^{n_{P,P'}} = g^{n'_k} \circ (g^{n^*_k})^{-1}. \tag{11.22}$$

From the estimate (3.27) for parabolic composition in Section 3.5, we have

$$C^{-1} \leqslant \frac{\operatorname{diam}(\gamma^*_k \cap P'_k)}{|P_k|, |P'|(\delta(Q_k, P'))^{-\frac{1}{2}}} \leqslant C. \tag{11.23}$$

We also have, from the estimates on simple composition

$$C^{-1} \leqslant \frac{\operatorname{diam} \omega^*(P, P', \varepsilon)|P^*_k|}{\operatorname{diam} \gamma^*_k} \leqslant C, \tag{11.24}$$

$$C^{-1} \leqslant \frac{|P_k|}{|P||P^*_k|} \leqslant C. \tag{11.25}$$

Multiplying these three inequalities yields the Lemma. □

11.5.4. From (11.19), we see that $\mathcal{E}(\omega^*)$ is a set of the type considered in Proposition 61. The property of uniformly bounded distortion has been checked in Lemma 22, and the property of uniform expansion is clear from Lemma 23.

Let us introduce

$$\chi(d) = \sum_{(P,P',\varepsilon)} [\operatorname{diam} \omega^*(P, P', \varepsilon)]^d. \tag{11.26}$$

If we are able, for some value of d, to show that the series defining χ is convergent and $\chi(d)$ is small enough, then by Proposition 61, we will deduce that the Hausdorff dimension of $\mathcal{E}(\omega^*)$ is $\leqslant d$.

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 23, we define, for $P \in C(\omega^*)$:

$$\chi_P(d) = \sum_{P'} |P'|^d \delta(\omega^*_P, P')^{-\frac{1}{2}d}. \tag{11.27}$$

We will then have

$$\chi(d) \leqslant C \sum_{P} |P|^d \chi_P(d). \tag{11.28}$$

In the sum (11.27), P′ is a transverse vertex of level > 0. We claim that

$$\delta(\omega^*_P, P') \leqslant \delta_{\max} := \min(\varepsilon_0, C|Q|^{1-\eta}). \tag{11.29}$$

The bound by ε_0 is clear. To show that $\delta(\omega^*_P, P') \leqslant C|Q|^{1-\eta}$, we recall that, for large k and small I containing t, Q_k and P′ are I-transverse, while Q and P′ are I-critically related for any I containing t. By Proposition 21 in Section 8.1, we must have $|Q| > \frac{1}{3}|P'|$. If we had $\delta(\omega^*_P, P') \gg |Q|^{1-\eta}$, we would also have $\delta(Q, P') \gg |Q|^{1-\eta}$, $\delta(Q, P') \gg |P'|^{1-\eta}$; conditions (T1), (T2), (T3) for $Q \,\overline{\pitchfork}_I\, P'$ would be satisfied for I small enough, a contradiction. The claim is proved.

In the series (11.27), we first sum over those P′ such that

$$2^{-\ell}\delta_{\max} \geqslant \delta(\omega^*_P, P') \geqslant 2^{-\ell-1}\delta_{\max} \tag{11.30}$$

for some fixed $\ell \geqslant 0$. This allows us to write

$$\chi_{\mathrm{P}}(d) \leqslant \mathrm{C}\delta_{\max}^{-d/2} \sum_{\ell \geqslant 0} 2^{\frac{\ell d}{2}} \Big(\sum^{(\ell)} |\mathrm{P}'|^d \Big), \tag{11.31}$$

where $\sum^{(\ell)}$ means that P′ is constrained by (11.30). We divide $\sum^{(\ell)}$ into two parts.

11.5.5. In the first part, denoted by $\sum_1^{(\ell)}$, we consider only those P′ such that its parent $\widetilde{\mathrm{P}}'$ satisfies

$$|\widetilde{\mathrm{P}}'| \leqslant 2^{-\ell}\delta_{\max}. \tag{11.32}$$

To estimate $\sum_1^{(\ell)} |\mathrm{P}'|^d$, first observe that, with d bounded away from 0, it follows from Proposition 26 in Section 8.1 that the sum of $|\mathrm{P}'|^d$ over children of a fixed parent $\widetilde{\mathrm{P}}'$ is bounded by $\mathrm{C}|\widetilde{\mathrm{P}}'|^d$. We must therefore bound $\sum_1^{(\ell)} |\widetilde{\mathrm{P}}'|^d$.

Also, as $\widetilde{\mathrm{P}}'$ is a critical vertex, $\widetilde{\mathrm{P}}'$ cannot be very thin: we have $\delta(\omega_{\mathrm{P}}^*, \mathrm{P}') \leqslant 3|\mathrm{P}'|^{1-\eta}$. Indeed, otherwise, for large k and small I containing t, conditions (T1), (T2), (T3) for $\mathrm{Q}_k \,\overline{\pitchfork}_{\mathrm{I}}\, \mathrm{P}'$ would be satisfied. We have therefore

$$|\widetilde{\mathrm{P}}'| \geqslant \mathrm{C}^{-1}[\delta(\omega_{\mathrm{P}}^*, \mathrm{P}')]^{(1-\eta)^{-1}} \geqslant \mathrm{C}^{-1}(\delta_{\max}2^{-\ell})^{(1-\eta)^{-1}}. \tag{11.33}$$

Finally, the number of $\widetilde{\mathrm{P}}'$ with $|\widetilde{\mathrm{P}}'|$ of order $2^{-m-\ell}\delta_{\max}$ is at most $\mathrm{C}2^m$ and the integer m here is restricted by (11.33) to the range

$$1 \leqslant 2^m \leqslant \mathrm{C}(\delta_{\max}2^{-\ell})^{-\eta(1-\eta)^{-1}}. \tag{11.34}$$

We, therefore, obtain for d bounded away from 0 and 1,

$$\begin{aligned} \sum_1^{(\ell)} |\mathrm{P}'|^d &\leqslant \mathrm{C} \sum_1^{(\ell)} |\widetilde{\mathrm{P}}'|^d \\ &\leqslant \mathrm{C}\delta_{\max}^d 2^{-\ell d} \sum_m 2^{m(1-d)} \\ &\leqslant \mathrm{C}(\delta_{\max}2^{-\ell})^{d-\eta}. \end{aligned} \tag{11.35}$$

11.5.6. In the second part of $\sum^{(\ell)}$, denoted by $\sum_2^{(\ell)}$, we have on the opposite

$$|\widetilde{\mathrm{P}}'| > 2^{-\ell}\delta_{\max}. \tag{11.36}$$

We claim that, in this case, the number of possible $\widetilde{\mathrm{P}}'$ is bounded by C.

– If there is some $\widetilde{P}'$ with $|\widetilde{P}'| \gg 2^{-\ell}\delta_{\max}$, it will contain all P′ satisfying (11.30), because (according to (MP7)) there is a strip of width $\geqslant C^{-1}|\widetilde{P}'|$ along each vertical boundary of $\widetilde{P}'$ which does not intersect Λ and therefore does not contain any P′ satisfying (11.30). Two such $\widetilde{P}'$ cannot both be the parents of the same P′, hence there is at most one such $\widetilde{P}'$.
– The remaining $\widetilde{P}'$ satisfy $C2^{-\ell}\delta_{\max} > |\widetilde{P}'| > 2^{-\ell}\delta_{\max}$. We divide this range into a bounded number of shorter range, such that two $\widetilde{P}'$ with widths in the same range are disjoint. But there can be at most C disjoint $\widetilde{P}'$ with $|\widetilde{P}'| > 2^{-\ell}\delta_{\max}$ and containing some P′ satisfying (11.30).

This proves the claim.

As each P′ is a transverse vertex, we must have (by (R7))

$$|P'| \leqslant C(2^{-\ell}\delta_{\max})^{(1-\eta)^{-1}}. \tag{11.37}$$

In particular, from (11.36), (11.37), P′ is a non-simple child of $\widetilde{P}'$. From Proposition 26 in Section 8.2, the number of P′ with |P′| of order $2^{-m}\varepsilon_0$ is at most $2^{c'm\eta}$.

We have

$$\sum_2^{(\ell)} |P'|^d \leqslant \varepsilon_0^d \sum_m 2^{-m(d-c'\eta)} \leqslant C\varepsilon_0^{c'\eta}(2^{-\ell}\delta_{\max})^{d-c'\eta}. \tag{11.38}$$

11.5.7. Putting (11.35) and (11.38) together yields (replacing if necessary c' by $\max(c', 1)$)

$$\sum^{(\ell)} |P'|^d \leqslant C(\delta_{\max}2^{-\ell})^{d-c'\eta} \tag{11.39}$$

and introducing this in (11.31) allows us to estimate χ_P:

$$\chi_P(d) \leqslant C\delta_{\max}^{\frac{1}{2}d-c'\eta}. \tag{11.40}$$

Finally, we obtain

$$\chi(d) \leqslant C \sum_{C(\omega^*)} |P|^d[\min(\varepsilon_0, |Q|)]^{\frac{1}{2}d-C\eta}. \tag{11.41}$$

11.5.8. 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

$$\frac{1}{p} + \frac{1}{q} = 1, \tag{11.42}$$

we have

$$\textbf{(11.43)} \qquad \chi(d) \leqslant C\chi_+(d)^{\frac{1}{p}} \chi_-(d)^{\frac{1}{q}}$$

where

$$\textbf{(11.44)} \qquad \chi_+(d) = \sum_{Q \text{ critical}} |P|^{dp},$$

$$\textbf{(11.45)} \qquad \chi_-(d) = \sum_{Q \text{ critical}} \min(\varepsilon_0, |Q|)^{(\frac{1}{2}d - C\eta)q}.$$

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 $\leqslant d$.

11.5.9. We first consider $\chi_-(d)$. We write $d^- = q(\frac{1}{2}d - C\eta)$. We divide the sum for $\chi_-(d)$ into three parts according to the width of Q. We recall that $|Q_u|$ is of the same order than ε_0.

- $|Q| \geqslant |Q_u|$.
  In this case, as Q is critical, it contains Q_u. The number of such Q is $\leqslant C \log \varepsilon_0^{-1}$, hence this part of the sum is bounded by $C\varepsilon_0^{d^-} \log \varepsilon_0^{-1}$ hence this part of the sum is bounded by $C\varepsilon_0^{d^-} \log \varepsilon_0^{-1}$.
- $|Q_u| > |Q| \geqslant C_0^{-1}\varepsilon_0^{1+\tau} = C_0^{-1}\varepsilon_1$.
  In this case, Q must be contained in Q_u. By Corollary 13 (Section 8.3), the number of such Q is at most $C\varepsilon_0^{-\tau d_u^*}$ and the corresponding part of the sum is bounded by $C\varepsilon_0^{d^- - \tau d_u^*}$.
- $C_0^{-1}\varepsilon_l > |Q| \geqslant C_0^{-1}\varepsilon_{l+1}$ for some $l > 0$.
  We first fix l and will afterwards sum over l.
  Let I be the parameter interval of length ε_{l-1} containing t. By Corollary 6 (Section 6.6.3), Q is I-defined. As Q is I-critical, it is I-special; by Proposition 18 (Section 7.7), we have $\max_I |Q| \leqslant \varepsilon_l$ (we choose C_0 in order to obtain this). By definition of $\mathcal{C}_-(I)$ (in Section 9.1), there exists $(P', Q', n') \in \mathcal{C}_-(I)$ such that $Q \subset Q'$. We have $\max_I |Q'| \leqslant \varepsilon_l$ and $|Q| \geqslant C_0^{-1}\varepsilon_{l+1}$; by Corollary 13, the number of Q for given Q' is $\leqslant C\varepsilon_l^{-\tau d_u^*}$. On the other hand, as I is strongly regular, we have from $(SR1)_u$ (Section 9.1) that

$$\textbf{(11.46)} \qquad \#\mathcal{C}_-(I) \leqslant C\left(\frac{|I|}{\varepsilon_0}\right)^{\sigma} \varepsilon_0^{-\tau d_u^0},$$

where σ, defined in (9.49) (Section 9.4), is close to $1 - d_u^0 - d_s^0$.

The part of the sum for $\chi_-(d)$ corresponding to this fixed value of l is therefore bounded by

$$C\varepsilon_0^{-\tau d_u^0}\left(\frac{\varepsilon_{l-1}}{\varepsilon_0}\right)^{\sigma}\varepsilon_l^{-\tau d_u^*}\varepsilon_l^{d^-} = C\varepsilon_0^{-\sigma-\tau d_u^0}\varepsilon_{l-1}^{\sigma+(1+\tau)(d^- -\tau d_u^*)}. \tag{11.47}$$

For the sum over l to converge, we ask that

$$d^- \geqslant -\frac{\sigma}{1+\tau}+\tau d_u^*+\tau, \tag{11.48}$$

where the right-hand side is close to $d_s^0+d_u^0-1$. When this is the case, the sum over l is, for ε_0 small enough, bounded by $C\varepsilon_0^{-\sigma-\tau d_u^0+\tau}$. From the definition of σ in (9.49), we have, as $1-d_s^0-d_u^0\leqslant 0$, that $\sigma\leqslant -\tau$.

We conclude that, if (11.48) holds, we have $\chi_-(d)<\varepsilon_0^{\tau}$.

11.5.10. We now consider $\chi_+(d)$. We write $d^+=pd$. We divide again the sum for $\chi_+(d)$ into three parts according to the width of Q.

– $|Q|\geqslant|Q_u|$.
  As Q must contain Q_u, this part of the sum is bounded by C (if d^+ is bounded away from 0, which will be the case).
– $|Q_u|>|Q|\geqslant(2\varepsilon_0)^{(1-\eta)^{-1}}$.
  In this case, Q must be contained in Q_u. By Corollary 13 (Section 8.3), the number of such Q is at most $C\varepsilon_0^{-C\eta}$. By Corollary 6 (Section 6.6.3), Q is I_0-defined, hence (P,Q,n) is the simple composition of some element in $\mathcal{R}(I_0)$ by (P_u,Q_u,n_u). In particular, we have $|P|\leqslant C|P_u|$ and this part of the sum is bounded by $\varepsilon_0^{-C\eta}|P_u|^{d^+}\leqslant C$ (with d^+ bounded away from 0).
– $(2\varepsilon_l)^{(1-\eta)^{-1}}>|Q|\geqslant(2\varepsilon_{l+1})^{(1-\eta)^{-1}}$ for some $l\geqslant 0$.
  We first fix l and will afterwards sum over l.
  Let I be the parameter interval of length ε_l containing t. By Corollary 6 (Section 6.6.3), Q is I-defined; as $(2\varepsilon_l)^{(1-\eta)^{-1}}>|Q|$, Q is thin I-critical. By definition of $\widehat{\mathcal{C}}_-(I)$ (Section 9.1), there exists $(P',Q',n')\in\widehat{\mathcal{C}}_-(I)$ such that $Q\subset Q'$. As I is strongly regular, condition $(SR2)'_u$ (Section 9.1) is satisfied.
  We take $d^+=\rho_s$, where ρ_s is the exponent in $(SR2)'_u$, defined in Section 9.14 and close to d_s^0. However, to apply $(SR2)'_u$, we need the Q to be disjoint. To achieve this, we divide the range

$$(2\varepsilon_l)^{(1-\eta)^{-1}}>|Q|\geqslant(2\varepsilon_{l+1})^{(1-\eta)^{-1}}$$

into several smaller ranges

$$(1-C^{-1})^i(2\varepsilon_l)^{(1-\eta)^{-1}}>|Q|\geqslant(1-C^{-1})^{i+1}(2\varepsilon_l)^{(1-\eta)^{-1}},\quad i=0,1,\dots$$

where C is large enough to insure that two Q with widths in the same range are disjoint. The number of smaller ranges is $\leqslant C\tau \log \varepsilon_l^{-1}$. The bound for $\sum |P|^{d^+}$ in each range given by $(SR2)'_u$ is $C|P_u|^{\rho_s}(\frac{\varepsilon_l}{\varepsilon_0})^{\sigma_s}$, where $\sigma_s = 1 - d_s^+ - 3\tau$ (cf. (9.63) in Section 9.6.1) is close to $1 - d_s^0$. Therefore the contribution to the sum for $\chi_+(d)$ coming for this value of l is bounded by

$$C\tau |P_u|^{\rho_s} \left(\frac{\varepsilon_l}{\varepsilon_0}\right)^{\sigma_s} \log \varepsilon_l^{-1}.$$

The sum over $l \geqslant 0$ is clearly converging and bounded by C (much better actually).

With $d^+ = \rho_s$, we obtain therefore $\chi_+(d) \leqslant C$.

11.5.11. We can now choose d, p, q to finish the proof. We take $\frac{1}{p} = \frac{d}{\rho_s}$, as already mentioned. In order to satisfy (11.48), we take

$$\frac{1}{q} = \frac{d - 2C\eta}{2(-\frac{\sigma}{1+\tau} + \tau d_u^* + \tau)}.$$

The relation $\frac{1}{p} + \frac{1}{q} = 1$ determines d. When $\tau \gg \eta \gg \varepsilon_0$ are small, we obtain a value for d close to

$$\left(\frac{1}{d_s^0} + \frac{1}{2(d_s^0 + d_u^0 - 1)}\right)^{-1} < d_s^0.$$

As explained above, the Hausdorff dimension of $\mathcal{E}(\omega^*)$ is $\leqslant d$ and the proof of the theorem is complete.

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 **6.** — *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:

$$W^s(\Lambda) = \bigcup_{n\geqslant 0} g^{-n}(W^s(\Lambda, \widehat{R}) \cap R). \tag{11.49}$$

Therefore, it is sufficient to show that $W^s(\Lambda, \widehat{R}) \cap R$ has Lebesgue measure 0. We write

$$R \cap W^s(\Lambda, \widehat{R}) = \bigcup_{n\geqslant 0} \Big(W^s(\Lambda, \widehat{R}) \cap R \cap g^{-n}(\widetilde{\mathcal{R}}_+^\infty)\Big) \sqcup \mathcal{E}^+, \tag{11.50}$$

with

$$(\mathbf{11.51})\qquad \mathcal{E}^+ = \{z \in W^s(\Lambda, \widehat{R}) \cap R, g^n(z) \notin \widetilde{\mathcal{R}}^\infty_+ \text{ for all } n \geqslant 0\}.$$

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

$$(\mathbf{11.52})\qquad \mathcal{E}^+ = \bigsqcup_{P_0} \mathcal{E}^+(P_0),$$

where

$$(\mathbf{11.53})\qquad \mathcal{E}^+(P_0) = \mathcal{E}^+ \cap c(P_0)$$

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

$$(\mathbf{11.54})\qquad g^{n_0}(\mathcal{E}^+(P_0)) \subset Q_0 \cap L_u \cap \mathcal{E}^+,$$

$$(\mathbf{11.55})\qquad g^{n_0+N_0}(\mathcal{E}^+(P_0)) \subset L_s \cap \mathcal{E}^+.$$

For $P_1 \in \mathcal{C}_-$, define

$$(\mathbf{11.56})\qquad \mathcal{E}^+(P_0, P_1) = \{z \in \mathcal{E}^+(P_0), g^{n_0+N_0}(z) \in c(P_1)\}.$$

We have a partition

$$(\mathbf{11.57})\qquad \mathcal{E}^+(P_0) = \bigsqcup_{P_1} \mathcal{E}^+(P_0, P_1).$$

At step k, we have a partition

$$(\mathbf{11.58})\qquad \mathcal{E}^+ = \bigsqcup_{P_0,\ldots,P_k} \mathcal{E}^+(P_0, \ldots, P_k)$$

where the (P_i, Q_i, n_i) run through $\mathcal{C}_-$. We write

$$m_0 = n_0,$$

$$m_1 = n_0 + N_0 + n_1,$$

$$(\mathbf{11.59}) \qquad \vdots$$

$$\begin{aligned} m_j &= n_0 + N_0 + n_1 + N_0 + \cdots + n_{j-1} + N_0 + n_j \\ &= m_{j-1} + N_0 + n_j. \end{aligned}$$

For $0 \leqslant j \leqslant k$, we have

$$(\mathbf{11.60}) \qquad g^{m_j}(\mathcal{E}^+(P_0, \ldots, P_k)) \subset Q_j \cap L_u \cap \mathcal{E}^+,$$

$$(\mathbf{11.61}) \qquad g^{m_j + N_0}(\mathcal{E}^+(P_0, \ldots, P_k)) \subset L_s \cap \mathcal{E}^+.$$

We define, for $P_{k+1} \in \mathcal{C}_-$

$$(\mathbf{11.62}) \qquad \mathcal{E}^+(P_0, \ldots, P_k, P_{k+1}) = \{z \in \mathcal{E}^+(P_0, \ldots, P_k), g^{m_k + N_0}(z) \in c(P_{k+1})\}$$

and we have

$$(\mathbf{11.63}) \qquad \mathcal{E}^+(P_0, \ldots, P_k) = \bigsqcup_{P_{k+1}} \mathcal{E}^+(P_0, \ldots, P_k, P_{k+1}).$$

However, in order to have $\mathcal{E}^+(P_0, \ldots, P_k) \neq \emptyset$ 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

$$(\mathbf{11.64}) \qquad \max(|P_1|, |Q_1|) \leqslant \varepsilon_0^{\beta}.$$

Lemma **24.** — *Let* $k \geqslant 1$. *Assume that* $\mathcal{E}^+(P_0, \ldots, P_{k+1})$ *is not-empty. We have*

$$\max(|P_{k+1}|, |Q_{k+1}|) \leqslant C|Q_k|^{\widetilde{\beta}},$$

with $\widetilde{\beta} = \beta(1-\eta)(1+\tau)^{-1}$.

Proof. — Let I be the largest parameter interval containing t such that

$$(\mathbf{11.65}) \qquad \max(|P_{k+1}|, |Q_{k+1}|) > |I|^{\beta}.$$

We first observe that, as $P_{k+1} \subset P_s$ and $Q_{k+1} \subset Q_u$, we have $\max(|P_{k+1}|, |Q_{k+1}|) \leqslant \varepsilon_0^{\beta}$ and I is not the starting interval I_0. Therefore, we have, by definition of I

$$(\mathbf{11.66}) \qquad \max(|P_{k+1}|, |Q_{k+1}|) \leqslant |I|^{\frac{\beta}{1+\tau}}.$$

Let $(\widetilde{P}_{k+1}, \widetilde{Q}_{k+1}, \widetilde{n}_{k+1}) \in \mathcal{R}(I)$ be the element such that $\widetilde{P}_{k+1}$ is the thinnest I-defined rectangle containing P_{k+1}. We claim that $\widetilde{P}_{k+1}$ is I-transverse. Indeed

– If $(\widetilde{P}_{k+1}, \widetilde{Q}_{k+1}, \widetilde{n}_{k+1}) = (P_{k+1}, Q_{k+1}, n_{k+1})$, this follows from (11.65), as Q_{k+1} is I-critical and I is β-regular.

– If $(\widetilde{P}_{k+1}, \widetilde{Q}_{k+1}, \widetilde{n}_{k+1}) \neq (P_{k+1}, Q_{k+1}, n_{k+1})$, from Corollary 5 (Section 6.6.3), we have that $|Q_{k+1}| \leqslant |I|^{\beta+\frac{1}{3}}$ and therefore $|\widetilde{P}_{k+1}| > |P_{k+1}| > |I|^{\beta}$. As $\widetilde{Q}_{k+1}$ is I-critical (by the structure theorem of Section 6.5), $\widetilde{P}_{k+1}$ must again be I-transverse.

The claim is proved. Next, we observe that Q_k and $\widetilde{P}_{k+1}$ cannot be I-separated, because $G(Q_k \cap L_u) \cap c(P_{k+1})$ contains $g^{m_k+N_0}(\mathcal{E}^+(P_0, \ldots, P_{k+1}))$. On the other hand, there cannot exist an element $(P^*, Q^*, n^*) \in \mathcal{R}(I)$ with $Q^* \supset Q_k$ and $Q^* \pitchfork_I P_{k+1}$: we would have $Q_k \pitchfork_{I'} P_{k+1}$ for small I′ and the corresponding non-simple descendants of P_k would contain $g^{m_{k-1}+N_0}(\mathcal{E}^+(P_0, \ldots, P_{k+1}))$, in contradiction with the definition of $c(P_k)$.

Therefore, the I-transversality of $\widetilde{P}_{k+1}$ implies the existence of $(P', Q', n') \in \mathcal{R}(I)$ with $Q' \subset Q_k$ and $Q' \pitchfork_I \widetilde{P}_{k+1}$. By coherence, Q_k is I-defined. By Proposition 10 (Section 6.4), as Q_k and $\widetilde{P}_{k+1}$ are not I-transverse, we must have $2|Q_k|^{1-\eta} > |I|$.

The estimate of the lemma follows from this and (11.66). □

Taking $\widehat{\beta} < \widetilde{\beta}$ but close to β and ε_0 sufficiently small, the estimate of the lemma and (11.64) give

$$\max(|P_j|, |Q_j|) \leqslant \varepsilon_0^{\widehat{\beta}^j}. \tag{11.67}$$

11.8. *Size and area of parabolic cores.*

Proposition **62.** — *Let* $(P, Q, n) \in \mathcal{C}_-$. *With* Leb *standing for Lebesgue measure, we have*

$$\operatorname{diam}(g^n(c(P))) \leqslant C|Q|^{\frac{1}{2}(1-\eta)}, \tag{11.68}$$

$$\operatorname{Leb}(g^n(c(P))) \leqslant C|Q|^{\frac{3}{2}-\frac{1}{2}\eta}, \tag{11.69}$$

$$\operatorname{Leb}(c(P)) \leqslant C|P||Q|^{\frac{1}{2}(1-\eta)}. \tag{11.70}$$

Remark **20.** — A posteriori, $c(P)$, which is contained in $W^s(\Lambda, \widehat{R})$, 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 Section 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 (11.70) in the proposition is a consequence of the second. On the other hand, as $g^n(c(P)) \subset Q$, the second inequality (11.69) is an obvious consequence of the first. We have, therefore, only to prove (11.68). Set

$$Z = g^{n+N_0}(c(P)). \tag{11.71}$$

This set is contained in $g^{N_0}(Q \cap L_u) \cap W^s(\Lambda, \widehat{R})$, and, a fortiori, in P_s. we have to show that diam $Z \leqslant C|Q|^{\frac{1}{2}(1-\eta)}$.

In order to do this, we will use the machinery introduced in the proof of Theorem 4, Sections 10.10.2 and 10.10.3. We extended the lamination of P_s by stable curves into a foliation by C^{1+Lip} vertical-like curves $\{x_s = \varphi(y_s, \overline{x})\}$ with Lipschitzian holonomy. Then, given a horizontal-like curve $\gamma = \{y_u = \psi(x_u\}$ in Q_u satisfying (10.155), we introduced a function $C(w, \overline{x})$ (vanishing at the points of intersection of $G_0 \circ G_-(\gamma)$ with $G_+(\{x_s = \varphi(y_s, \overline{x})\})$) and a function $\delta(\overline{x}) = -\min_w C(w, \overline{x})$, whose properties are given in Lemma 19 in Section 10.10.3.

We choose for γ one of the two horizontal-like boundary curves of Q, more precisely the one which gives the greater values of δ (corresponding to δ_L, δ_{LR} in the context of Section 3.6.

Lemma **25.** — *If a curve* $\{x_s = \varphi(y_s, \overline{x})\}$ *contains a point of* Z, *then we have*

$$0 \leqslant \delta(\overline{x}) \leqslant C|Q|^{1-\eta}.$$

Proof. — That $\delta(\overline{x}) \geqslant 0$ follows from the choice of γ and the fact that $Z \subset G(Q \cap L_u)$. We prove the other inequality by contradiction, assuming that there is a point $z' = g^{n+N_0}(z) = (\varphi(y_s, \overline{x}), y_s) \in Z$ with $\delta(\overline{x}) \geqslant C_0|Q|^{1-\eta}$, C_0 large.

Let $(P', Q', n') \in \mathcal{R}$ be an element such that $z' \in P'$. We claim that $|P'| \geqslant 3|Q'|$. Indeed, if we had $|P'| < 3|Q'|$, the inequality $\delta(\overline{x}) \geqslant C_0|Q|^{1-\eta}$ with C_0 large enough would directly imply (T1), (T2), (T3) for $Q \overline{\pitchfork}_I P'$, I small enough. But then z would belong to a descendant of P, in contradiction with the definition of $c(P)$. The claim is proved.

It follows from the claim that there exists a thinnest rectangle P' containing z'. As z' belongs to $W_s(\Lambda)$ but not to any child of P', P' is not I-decomposable for any parameter interval I containing t; therefore P' is I-transverse for I small enough.

As $z' \in G(Q \cap L_u) \cap P'$, Q and P' are not I-separated. We have also already seen that they cannot be I-transverse. Therefore, as P' is I-transverse, there exists $(P^*, Q^*, n^*) \in \mathcal{R}$ with $Q \supset Q^*$ such that Q^* and P' are I-transverse. But then, by Proposition 21 (Section 8.1), it follows from $|P'| \geqslant 3|Q'|$ that $Q \pitchfork_{I'} P'$ holds for I' small enough, a contradiction. □

We have shown that Z is contained in the lenticular region bounded on one side by $G(\gamma \cap L_u)$ and on the other by the curve $\{x_s = \varphi(y_s, \overline{x}^*)\}$, with $\delta(\overline{x}^*) = C|Q|^{1-\eta}$. The quadratic geometry of C, δ given by Lemma 19 guarantees that this lenticular region has diameter $\leqslant C|Q|^{\frac{1}{2}(1-\eta)}$. □

11.9. *Proof of Theorem 6.* — We will estimate first the Lebesgue measure of each domain $\mathcal{E}^+(P_0, \dots, P_k)$. We have

$$g^{m_{k-1}+N_0}(\mathcal{E}^+(P_0, \dots, P_k)) \subset c(P_k). \tag{11.72}$$

We now use that both the fixed map g^{N_0} and the affine-like iterates $g^{n_j} : P_j \to Q_j$ (for $0 \leqslant j < k$) have uniformly bounded distortion with respect to Lebesgue measure. We are, therefore, able to deduce from (11.69) in Proposition 62 that

$$\text{(11.73)} \qquad \text{Leb}(\mathcal{E}^+(P_0, \dots, P_k)) \leqslant C^{k+1} |Q_k|^{\frac{3}{2} - \frac{1}{2}\eta} \prod_0^k \frac{|P_j|}{|Q_j|}.$$

By Lemma 24, 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 $|Q_0| \geqslant \varepsilon_0$; if $Q_0 \subset Q_u$, the argument of Lemma 24 applies). It then follows from (11.73) that we have (for $k > 0$)

$$\text{(11.74)} \qquad \text{Leb}(\mathcal{E}^+(P_0, \dots, P_k)) \ll |P_0| |Q_k|^{\frac{1}{2}(1-\eta)}.$$

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

$$\text{(11.75)} \qquad |Q_{k-1}| \geqslant C^{-1} \max(|P_k| |Q_k|)^{1/\widetilde{\beta}}.$$

On the other hand, as Q_{k-1}, P_k are neither separated nor transverse (for every parameter interval containing t), we must have, as $|Q_{k-1}| \gg |P_k|$

$$\text{(11.76)} \qquad 0 \leqslant \delta_{LR}(Q_{k-1}, P_k) \leqslant C |Q_{k-1}|^{1-\eta}.$$

For every scale $2^{-l} \geqslant C^{-1} \max(|P_k||Q_k|)^{1/\widetilde{\beta}}$, this will give at most $C 2^{l\eta}$ possibilities for Q_{k-1} with $2^{-l} \geqslant |Q_{k-1}| > 2^{-l-1}$. Summing over l gives at most $C |Q_k|^{-\eta/\widetilde{\beta}}$ total possibilities for Q_{k-1}.

We repeat this, with Q_{k-1} now fixed, for $Q_{k-2}, \dots$. We obtain a number of possible admissible sequence not greater than

$$\text{(11.77)} \qquad C^l |Q_k|^{-\eta(\widetilde{\beta}^{-1} + \dots + \widetilde{\beta}^{-l})}$$

with $l \leqslant C \log \log |Q_k|^{-1}$. We conclude that the total number of admissible sequences with fixed extremity Q_k is bounded by $C |Q_k|^{-C\eta}$.

Therefore, we obtain

$$\text{(11.78)} \qquad \text{Leb}(\mathcal{E}^+) \leqslant \sum_{P_0, Q_k} |P_0| |Q_k|^{\frac{1}{2} - C\eta}.$$

We first claim that the sum $\sum |P_0|$ is bounded. Actually, this is true even if we do not restrict $(P_0, Q_0, n_0) \in \mathcal{R}$ by the condition that Q_0 is critical. Indeed, when we sum over a given generation (with a fixed number of ascendants), the P_0 are disjoint hence the sum is bounded. But the sum of the widths of the children is always smaller by a definite factor than the width of the parent; therefore, the sum over the various generations is bounded.

Finally, we claim that the sum $\sum_{Q_k} |Q_k|^{\frac{1}{2}-C\eta}$ is arbitrarily small (when k is large). Indeed, it has been shown in Section 11.5.9 that the sum $\sum_{Q \text{ critical}} |Q|^{d^-}$ is convergent provided (cf. (11.48))

$$(\mathbf{11.79}) \qquad d^- \geqslant -\frac{\sigma}{1+\tau} + \tau d_u^* + \tau.$$

Here, the right-hand side is close to $d_s^0 + d_u^0 - 1$ (when $\tau \gg \eta \gg \varepsilon_0$ are small). On the other hand, under condition (H4), the maximum value of $d_s^0 + d_u^0 - 1$ is $1/5$. Therefore, the sum $\sum_{Q \text{ critical}} |Q|^{\frac{1}{2}-C\eta}$ is convergent. But we have

$$(\mathbf{11.80}) \qquad |Q_k| \leqslant \varepsilon_0^{\widehat{\beta}^k}.$$

We deduce that

$$(\mathbf{11.81}) \qquad \lim_{k\to+\infty} \sum_{Q_k} |Q_k|^{\frac{1}{2}-C\eta} = 0,$$

and this concludes the proof of Theorem 6.

We can sum up the results in Sections 10 and 11 by rephrasing our main result as follows:

Theorem **7.** — *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 repellors.*

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 *Formulas for simple composition.* — Here we consider a map $F_t : (x_0, y_0) \mapsto (x_1, y_1)$ implicitly defined by

$$(\mathbf{A.1}) \qquad \begin{cases} x_0 = A(y_0, x_1, t), \\ y_1 = B(y_0, x_1, t), \end{cases}$$

and a map $F'_t : (x_1, y_1) \mapsto (x_2, y_2)$ implicitly defined by

$$\begin{cases} x_1 = A'(y_1, x_2, t), \\ y_2 = B'(y_1, x_2, t). \end{cases} \tag{A.2}$$

The composition $F''_t = F'_t \circ F_t$ is implicitly defined by

$$\begin{cases} x_0 = A''(y_0, x_2, t), \\ y_2 = B''(y_0, x_2, t) \end{cases} \tag{A.3}$$

and we want to relate the partial derivatives of A'', B'' to those of A, B, A', B'. Set

$$\Delta := 1 - A'_y(y_1, x_2, t)B_x(y_0, x_1, t). \tag{A.4}$$

When we solve the system (A.1), (A.2) for x_1, y_1, we obtain

$$\begin{cases} x_1 = X(y_0, x_2, t), \\ y_1 = Y(y_0, x_2, t), \end{cases} \tag{A.5}$$

where the partial derivatives of X, Y are given by

$$\begin{cases} X_x = A'_x \Delta^{-1}, \\ X_y = A'_y B_y \Delta^{-1}, \\ X_t = (A'_t + A'_y B_t)\Delta^{-1}, \\ Y_x = A'_x B_x \Delta^{-1}, \\ Y_y = B_y \Delta^{-1}, \\ Y_t = (B_t + A'_t B_x)\Delta^{-1}. \end{cases} \tag{A.6}$$

We have

$$\begin{cases} A''(y_0, x_2, t) = A(y_0, X, t), \\ B''(y_0, x_2, t) = B'(Y, x_2, t), \end{cases} \tag{A.7}$$

which gives

$$\begin{cases} A''_x = A_x A'_x \Delta^{-1}, \\ B''_y = B'_y B_y \Delta^{-1}, \end{cases} \tag{A.8}$$

$$\begin{cases} A''_y = A_y + A_x X_y, \\ B''_x = B'_x + B'_y Y_x, \end{cases} \tag{A.9}$$

$$\textbf{(A.10)}\qquad \begin{cases} A''_t = A_t + A_x X_t, \\ B''_t = B'_t + B'_y Y_t. \end{cases}$$

Next, from (A.4), we have

$$\textbf{(A.11)}\qquad \begin{cases} -\Delta_x = B_{xx} X_x A'_y + B_x A'_{xy} + B_x A'_{yy} Y_x, \\ -\Delta_y = A'_{yy} Y_y B_x + A'_y B_{xy} + A'_y B_{xx} X_y, \\ -\Delta_t = B_{xt} A'_y + B_{xx} X_t A'_y + B_x A'_{yt} + B_x A'_{yy} Y_t. \end{cases}$$

Taking logarithmic derivatives in (A.8) gives

$$\textbf{(A.12)}\qquad \partial_x \log|A''_x| = \partial_x \log|A'_x| + Y_x \partial_y \log|A'_x| + X_x \partial_x \log|A_x| - \Delta_x \Delta^{-1},$$

$$\textbf{(A.13)}\qquad \partial_y \log|A''_x| = \partial_y \log|A_x| + X_y \partial_x \log|A_x| + Y_y \partial_y \log|A'_x| - \Delta_y \Delta^{-1},$$

$$\textbf{(A.14)}\qquad \begin{aligned} \partial_t \log|A''_x| = {} & \partial_t \log|A_x| + \partial_t \log|A'_x| + X_t \partial_x \log|A_x| \\ & + Y_t \partial_y \log|A'_x| - \Delta_t \Delta^{-1}, \end{aligned}$$

$$\textbf{(A.15)}\qquad \partial_y \log|B''_y| = \partial_y \log|B_y| + X_y \partial_x \log|B_y| + Y_y \partial_y \log|B'_y| - \Delta_y \Delta^{-1},$$

$$\textbf{(A.16)}\qquad \partial_x \log|B''_y| = \partial_x \log|B'_y| + Y_x \partial_y \log|B'_y| + X_x \partial_x \log|B_y| - \Delta_x \Delta^{-1},$$

$$\textbf{(A.17)}\qquad \begin{aligned} \partial_t \log|B''_y| = {} & \partial_t \log|B'_y| + \partial_t \log|B_y| + Y_t \partial_y \log|B'_y| \\ & + X_t \partial_x \log|B_y| - \Delta_t \Delta^{-1}. \end{aligned}$$

Taking derivatives in (A.9) gives

$$\textbf{(A.18)}\qquad \begin{cases} A''_{yy} = A_{yy} + 2A_{xy} X_y + A_{xx} X_y^2 + A_x X_{yy}, \\ B''_{xx} = B'_{xx} + 2B'_{xy} Y_x + B'_{yy} Y_x^2 + B'_y Y_{xx}, \end{cases}$$

$$\textbf{(A.19)}\qquad \begin{cases} A''_{yt} = A_{yt} + X_t A_{xy} + X_y A_{xt} + X_t X_y A_{xx} + A_x X_{yt}, \\ B''_{xt} = B'_{xt} + Y_t B'_{xy} + Y_x B'_{yt} + Y_t Y_x B'_{yy} + B'_y Y_{xt}, \end{cases}$$

where the partial derivatives of X, Y are obtained from (A.6):

$$\textbf{(A.20)}\qquad X_{yy} = B_y \Delta^{-1} (A'_{yy} Y_y + A'_y \partial_y \log|B_y| + A'_y X_y \partial_x \log|B_y| - A'_y \Delta_y \Delta^{-1}),$$

$$\textbf{(A.21)}\qquad Y_{xx} = A'_x \Delta^{-1} (B_{xx} X_x + B_x \partial_x \log|A'_x| + B_x Y_x \partial_y \log|A'_x| - B_x \Delta_x \Delta^{-1}),$$

$$\textbf{(A.22)}\qquad X_{yt} = B_y \Delta^{-1} (A'_{yy} Y_t + A'_{yt} + A'_y \partial_t \log|B_y| + A'_y X_t \partial_x \log|B_y| - A'_y \Delta_t \Delta^{-1}),$$

$$\textbf{(A.23)}\qquad Y_{xt} = A'_x \Delta^{-1} (B_{xx} X_t + B_{xt} + B_x \partial_t \log|A'_x| + B_x Y_t \partial_y \log|A'_x| - B_x \Delta_t \Delta^{-1}).$$

A.2 *Formulas for 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

$$\begin{cases} y_s = Y_s(w, x_s, t), \\ x_u = X_u(w, y_u, t), \end{cases} \tag{A.24}$$

$$w^2 = \theta(y_u, x_s, t). \tag{A.25}$$

We also have an affine like map $F_0 : (x_0, y_0) \mapsto (x_u, y_u)$ implicitly defined by

$$\begin{cases} x_0 = A_0(y_0, x_u, t), \\ y_u = B_0(y_0, x_u, t), \end{cases} \tag{A.26}$$

and another affine-like map $F_1 : (x_1, y_1) \mapsto (x_1, y_1)$ implicitly defined by

$$\begin{cases} x_s = A_1(y_s, x_1, t), \\ y_1 = B_1(y_s, x_1, t). \end{cases} \tag{A.27}$$

We assume that (PC1), (PC2) in Section 3.5 are satisfied. As we have seen in [PY2] and Section 3.5, the first step is to write

$$\begin{cases} x_u = X(w, y_0, t), \\ y_s = Y(w, x_1, t), \end{cases} \tag{A.28}$$

where the partial derivatives of X, Y are given by

$$\begin{cases} X_w = X_{u,w} \Delta_0^{-1}, \\ X_y = X_{u,y} B_{0,y} \Delta_0^{-1}, \\ X_t = (X_{u,t} + X_{u,y} B_{0,t}) \Delta_0^{-1}, \\ Y_w = Y_{s,w} \Delta_1^{-1}, \\ Y_x = Y_{s,x} A_{1,x} \Delta_1^{-1}, \\ Y_t = (Y_{s,t} + Y_{s,x} A_{1,t}) \Delta_1^{-1}, \end{cases} \tag{A.29}$$

$$\begin{cases} \Delta_0 = 1 - X_{u,y} B_{0,x}, \\ \Delta_1 = 1 - Y_{s,x} A_{1,y}. \end{cases} \tag{A.30}$$

We set

$$\text{(A.31)} \qquad \begin{cases} \overline{Y}(w, y_0, t) := B_0(y_0, X, t), \\ \overline{X}(w, x_1, t) := A_1(Y, x_1, t), \end{cases}$$

$$\text{(A.32)} \qquad C(w, y_0, x_1, t) := w^2 - \theta(\overline{X}, \overline{Y}, t).$$

The partial derivatives are given by

$$\text{(A.33)} \qquad \begin{cases} \overline{Y}_w = B_{0,x} X_w, \\ \overline{Y}_y = B_{0,y} + B_{0,x} X_y = B_{0,y} \Delta_0^{-1}, \\ \overline{Y}_t = B_{0,t} + B_{0,x} X_t = (B_{0,t} + B_{0,x} X_{u,t}) \Delta_0^{-1}, \end{cases}$$

$$\text{(A.34)} \qquad \begin{cases} \overline{X}_w = A_{1,y} Y_w, \\ \overline{X}_x = A_{1,x} + A_{1,y} Y_x = A_{1,x} \Delta_1^{-1}, \\ \overline{X}_t = A_{1,t} + A_{1,y} Y_t = (A_{1,t} + A_{1,y} Y_{s,t}) \Delta_1^{-1}, \end{cases}$$

$$\text{(A.35)} \qquad \begin{cases} -C_w = -2w + \theta_x \overline{X}_w + \theta_y \overline{Y}_w, \\ -C_x = \theta_x \overline{X}_x, \\ -C_y = \theta_y \overline{Y}_y, \\ -C_t = \theta_x \overline{X}_t + \theta_y \overline{Y}_t + \theta_t. \end{cases}$$

We solve

$$\text{(A.36)} \qquad C(w, y_0, x_1, t) = 0$$

to define

$$\text{(A.37)} \qquad w = W(y_0, x_1, t)$$

(there are two solutions W^+ and W^-).

The corresponding branch of the parabolic composition is implicitly defined by

$$\text{(A.38)} \qquad \begin{cases} x_0 = A_0(y_0, X(W, y_0, t), t) =: A(y_0, x_1, t), \\ y_1 = B_1(Y(W, x_1, t), x_1, t) =: B(y_0, x_1, t). \end{cases}$$

The partial derivatives of A, B, W are given by

$$\text{(A.39)} \qquad \begin{cases} A_x = A_{0,x} X_w W_x, \\ A_y = A_{0,y} + A_{0,x}(X_y + X_w W_y), \\ A_t = A_{0,t} + A_{0,x}(X_t + X_w W_t), \end{cases}$$

$$\text{(A.40)}\qquad \begin{cases} B_y = B_{1,y} Y_w W_y, \\ B_x = B_{1,x} + B_{1,y}(Y_x + Y_w W_x), \\ B_t = B_{1,t} + B_{1,y}(Y_t + Y_w W_t), \end{cases}$$

$$\text{(A.41)}\qquad \begin{cases} W_x = -C_x C_w^{-1}, \\ W_y = -C_y C_w^{-1}, \\ W_t = -C_t C_w^{-1}. \end{cases}$$

Substituting (A.29), (A.41), (A.35), (A.34) in the formulas (A.39)–(A.40) leads to

$$\text{(A.42)}\qquad \begin{cases} A_x = A_{0,x} A_{1,x} C_w^{-1} \theta_x X_{u,w} \Delta_0^{-1} \Delta_1^{-1}, \\ B_y = B_{1,y} B_{0,y} C_w^{-1} \theta_y Y_{s,w} \Delta_0^{-1} \Delta_1^{-1}, \end{cases}$$

$$\text{(A.43)}\qquad \begin{cases} A_y = A_{0,y} + A_{0,x} B_{0,y} \Delta_0^{-1}(X_{u,y} + X_{u,w}\theta_y \Delta_0^{-1} C_w^{-1}), \\ B_x = B_{1,x} + B_{1,y} A_{1,x} \Delta_1^{-1}(Y_{s,x} + Y_{s,w}\theta_x \Delta_1^{-1} C_w^{-1}), \end{cases}$$

$$\text{(A.44)}\qquad \begin{cases} A_t = A_{0,t} + A_{0,x}\Delta_0^{-1}[X_{u,t} + X_{u,y} B_{0,t} + X_{u,w} C_w^{-1}(\theta_t + \theta_x \overline{X}_t + \theta_y \overline{Y}_t)], \\ B_t = B_{1,t} + B_{1,y}\Delta_1^{-1}[Y_{s,t} + Y_{s,x} A_{1,t} + Y_{s,w} C_w^{-1}(\theta_t + \theta_x \overline{X}_t + \theta_y \overline{Y}_t)]. \end{cases}$$

Taking the logarithmic derivatives in the first formula of (A.39), we obtain

$$\text{(A.45)}\qquad \partial_x \log|A_x| = W_x X_w \partial_x \log|A_{0,x}| + W_x \partial_w \log|X_w| + \partial_x \log|W_x|,$$

$$\begin{aligned} \partial_y \log|A_x| &= \partial_y \log|A_{0,x}| + \partial_x \log|A_{0,x}|(X_y + X_w W_y) \\ &\quad + \partial_y \log|X_w| + W_y \partial_w \log|X_w| + \partial_y \log|W_x|, \end{aligned} \qquad \text{(A.46)}$$

$$\begin{aligned} \partial_t \log|A_x| &= \partial_t \log|A_{0,x}| + \partial_x \log|A_{0,x}|(X_t + X_w W_t) \\ &\quad + \partial_t \log|X_w| + W_t \partial_w \log|X_w| + \partial_t \log|W_x|. \end{aligned} \qquad \text{(A.47)}$$

From the second formula in (A.39), one gets

$$\begin{aligned} A_{yy} &= A_{0,yy} + 2A_{0,xy}(X_y + X_w W_y) + A_{0,xx}(X_y + X_w W_y)^2 \\ &\quad + A_{0,x}(X_{yy} + 2X_{wy} W_y + X_{ww} W_y^2 + X_w W_{yy}), \end{aligned} \qquad \text{(A.48)}$$

$$\begin{aligned} A_{yt} &= A_{0,yt} + A_{0,xy}(X_y + X_w W_t) + A_{0,xt}(X_y + X_w W_y) \\ &\quad + A_{0,xx}(X_t + X_w W_t)(X_y + X_w W_y) \\ &\quad + A_{0,x}(X_{yt} + X_{wy} W_t + X_{wt} W_y + X_{ww} W_y W_t + X_w W_{yt}). \end{aligned} \qquad \text{(A.49)}$$

The symmetric formulas for B are

$$\text{(A.50)}\qquad \partial_y \log|B_y| = W_y Y_w \partial_y \log|B_{1,y}| + \partial_y \log|W_y| + W_y \partial_w \log|Y_w|,$$

$$\begin{aligned}\partial_x \log|\mathrm{B}_y| &= \partial_x \log|\mathrm{B}_{1,y}| + \partial_x \log|\mathrm{Y}_w| \\ &\quad + \partial_x \log|\mathrm{W}_y| + \partial_y \log|\mathrm{B}_{1,y}|(\mathrm{Y}_x + \mathrm{Y}_w \mathrm{W}_x) + \mathrm{W}_x \partial_w \log|\mathrm{Y}_w|,\end{aligned} \tag{A.51}$$

$$\begin{aligned}\partial_t \log|\mathrm{B}_y| &= \partial_t \log|\mathrm{B}_{1,y}| + \partial_y \log|\mathrm{B}_{1,y}|(\mathrm{Y}_t + \mathrm{Y}_w \mathrm{W}_t) \\ &\quad + \partial_t \log|\mathrm{Y}_w| + \mathrm{W}_t \partial_w \log|\mathrm{Y}_w| + \partial_t \log|\mathrm{W}_y|,\end{aligned} \tag{A.52}$$

$$\begin{aligned}\mathrm{B}_{xx} &= \mathrm{B}_{1,xx} + 2\mathrm{B}_{1,xy}(\mathrm{Y}_x + \mathrm{Y}_w \mathrm{W}_x) + \mathrm{B}_{1,yy}(\mathrm{Y}_x + \mathrm{Y}_w \mathrm{W}_x)^2 \\ &\quad + \mathrm{B}_{1,y}(\mathrm{Y}_{xx} + 2\mathrm{Y}_{wx}\mathrm{W}_x + \mathrm{Y}_{ww}\mathrm{W}_x^2 + \mathrm{Y}_w \mathrm{W}_{xx}),\end{aligned} \tag{A.53}$$

$$\begin{aligned}\mathrm{B}_{xt} &= \mathrm{B}_{1,xt} + \mathrm{B}_{1,xy}(\mathrm{Y}_t + \mathrm{Y}_w \mathrm{W}_t) + \mathrm{B}_{1,yt}(\mathrm{Y}_x + \mathrm{Y}_w \mathrm{W}_x) \\ &\quad + \mathrm{B}_{1,yy}(\mathrm{Y}_x + \mathrm{Y}_w \mathrm{W}_x)(\mathrm{Y}_t + \mathrm{Y}_w \mathrm{W}_t) \\ &\quad + \mathrm{B}_{1,y}(\mathrm{Y}_{xt} + \mathrm{Y}_{wx}\mathrm{W}_t + \mathrm{Y}_{wt}\mathrm{W}_x + \mathrm{Y}_{ww}\mathrm{W}_x\mathrm{W}_t + \mathrm{Y}_w \mathrm{W}_{xt}).\end{aligned} \tag{A.54}$$

In formulas (A.45)–(A.54), the partial derivatives of order 2 of W are obtained from (A.41):

$$\begin{cases}\mathrm{W}_{xx} = -\mathrm{C}_w^{-1}(\mathrm{C}_{ww}\mathrm{W}_x^2 + 2\mathrm{C}_{wx}\mathrm{W}_x + \mathrm{C}_{xx}),\\ \mathrm{W}_{xy} = -\mathrm{C}_w^{-1}(\mathrm{C}_{ww}\mathrm{W}_x\mathrm{W}_y + \mathrm{C}_{wx}\mathrm{W}_y + \mathrm{C}_{wy}\mathrm{W}_x + \mathrm{C}_{xy}),\\ \mathrm{W}_{yy} = -\mathrm{C}_w^{-1}(\mathrm{C}_{ww}\mathrm{W}_y^2 + 2\mathrm{C}_{wy}\mathrm{W}_y + \mathrm{C}_{yy}),\\ \mathrm{W}_{xt} = -\mathrm{C}_w^{-1}(\mathrm{C}_{ww}\mathrm{W}_x\mathrm{W}_t + \mathrm{C}_{wx}\mathrm{W}_t + \mathrm{C}_{wt}\mathrm{W}_x + \mathrm{C}_{xt}),\\ \mathrm{W}_{yt} = -\mathrm{C}_w^{-1}(\mathrm{C}_{ww}\mathrm{W}_y\mathrm{W}_t + \mathrm{C}_{wy}\mathrm{W}_t + \mathrm{C}_{wt}\mathrm{W}_y + \mathrm{C}_{yt}).\end{cases} \tag{A.55}$$

The partial derivatives of order 2 of C are obtained from (A.35):

$$-\mathrm{C}_{ww} = -2 + \theta_x \overline{\mathrm{X}}_{ww} + \theta_y \overline{\mathrm{Y}}_{ww} + \theta_{xx}\overline{\mathrm{X}}_w^2 + 2\theta_{xy}\overline{\mathrm{X}}_w\overline{\mathrm{Y}}_w + \theta_{yy}\overline{\mathrm{Y}}_w^2, \tag{A.56}$$

$$\begin{cases}-\mathrm{C}_{wx} = \theta_{xx}\overline{\mathrm{X}}_w\overline{\mathrm{X}}_x + \theta_{xy}\overline{\mathrm{Y}}_w\overline{\mathrm{X}}_x + \theta_x\overline{\mathrm{X}}_{wx},\\ -\mathrm{C}_{wy} = \theta_{yy}\overline{\mathrm{Y}}_w\overline{\mathrm{Y}}_y + \theta_{xy}\overline{\mathrm{X}}_w\overline{\mathrm{Y}}_y + \theta_y\overline{\mathrm{Y}}_{wy},\end{cases} \tag{A.57}$$

$$\begin{aligned}-\mathrm{C}_{wt} &= \theta_{xt}\overline{\mathrm{X}}_w + \theta_{xx}\overline{\mathrm{X}}_w\overline{\mathrm{X}}_t + \theta_{xy}(\overline{\mathrm{X}}_w\overline{\mathrm{Y}}_t + \overline{\mathrm{Y}}_w\overline{\mathrm{X}}_t) + \theta_{yy}\overline{\mathrm{Y}}_w\overline{\mathrm{Y}}_t \\ &\quad + \theta_{yt}\overline{\mathrm{Y}}_w + \theta_x\overline{\mathrm{X}}_{wt} + \theta_y\overline{\mathrm{Y}}_{wt},\end{aligned} \tag{A.58}$$

$$\begin{cases}-\mathrm{C}_{xx} = \theta_{xx}\overline{\mathrm{X}}_x^2 + \theta_x\overline{\mathrm{X}}_{xx},\\ -\mathrm{C}_{xy} = \theta_{xy}\overline{\mathrm{X}}_x\overline{\mathrm{Y}}_y,\\ -\mathrm{C}_{yy} = \theta_{yy}\overline{\mathrm{Y}}_y^2 + \theta_y\overline{\mathrm{Y}}_{yy},\end{cases} \tag{A.59}$$

$$\begin{cases}-\mathrm{C}_{xt} = \theta_{xt}\overline{\mathrm{X}}_x + \theta_{xx}\overline{\mathrm{X}}_x\overline{\mathrm{X}}_t + \theta_{xy}\overline{\mathrm{X}}_x\overline{\mathrm{Y}}_t + \theta_x\overline{\mathrm{X}}_{xt},\\ -\mathrm{C}_{yt} = \theta_{yt}\overline{\mathrm{Y}}_y + \theta_{yy}\overline{\mathrm{Y}}_y\overline{\mathrm{Y}}_t + \theta_{xy}\overline{\mathrm{Y}}_y\overline{\mathrm{X}}_t + \theta_y\overline{\mathrm{Y}}_{yt}.\end{cases} \tag{A.60}$$

The partial derivatives of order 2 of $\overline{X}$, $\overline{Y}$ are obtained from (A.34):

$$\text{(A.61)}\qquad \begin{cases} \overline{X}_{ww} = A_{1,yy}Y_w^2 + A_{1,y}Y_{ww}, \\ \overline{X}_{wx} = A_{1,xy}Y_w + A_{1,yy}Y_wY_x + A_{1,y}Y_{wx}, \\ \overline{X}_{wt} = A_{1,yt}Y_w + A_{1,yy}Y_wY_t + A_{1,y}Y_{wt}, \\ \overline{X}_{xx} = A_{1,xx} + 2A_{1,xy}Y_x + A_{1,yy}Y_x^2 + A_{1,y}Y_{xx}, \\ \overline{X}_{xt} = A_{1,xt} + A_{1,xy}Y_t + A_{1,yt}Y_x + A_{1,yy}Y_xY_t + A_{1,y}Y_{xt}, \end{cases}$$

$$\text{(A.62)}\qquad \begin{cases} \overline{Y}_{ww} = B_{0,xx}X_w^2 + B_{0,x}X_{ww}, \\ \overline{Y}_{wy} = B_{0,xy}X_w + B_{0,xx}X_wX_y + B_{0,x}X_{wy}, \\ \overline{Y}_{wt} = B_{0,xt}X_w + B_{0,xx}X_wX_t + B_{0,x}X_{wt}, \\ \overline{Y}_{yy} = B_{0,yy} + 2B_{0,xy}X_y + B_{0,xx}X_y^2 + B_{0,x}X_{yy}, \\ \overline{Y}_{yt} = B_{0,yt} + B_{0,xy}X_t + B_{0,xt}X_y + B_{0,xx}X_yX_t + B_{0,x}X_{yt}. \end{cases}$$

Finally, from (A.29), we obtain

$$\text{(A.63)}\qquad \begin{cases} X_{ww} = \Delta_0^{-1}(X_{u,ww} + 2X_{u,wy}\overline{Y}_w + X_{u,yy}\overline{Y}_w^2 + X_{u,y}B_{0,xx}X_w^2), \\ X_{wy} = \Delta_0^{-1}(X_{u,wy}\overline{Y}_y + X_{u,yy}\overline{Y}_w\overline{Y}_y + X_{u,y}X_w(B_{0,xy} + B_{0,xx}X_y)), \\ X_{yy} = \Delta_0^{-1}(X_{u,yy}\overline{Y}_y^2 + X_{u,y}(B_{0,yy} + 2B_{0,xy}X_y + B_{0,xx}X_y^2)), \\ X_{wt} = \Delta_0^{-1}[X_{u,wt} + X_{u,wy}\overline{Y}_t + X_{u,yy}\overline{Y}_w\overline{Y}_t + X_{u,yt}\overline{Y}_w \\ \qquad + X_{u,y}X_w(B_{0,xt} + B_{0,xx}X_t)], \\ X_{yt} = \Delta_0^{-1}[X_{u,yt}\overline{Y}_y + X_{u,yy}\overline{Y}_y\overline{Y}_t \\ \qquad + X_{u,y}(B_{0,yt} + B_{0,xy}X_t + B_{0,xt}X_y + B_{0,xx}X_yX_t)], \end{cases}$$

$$\text{(A.64)}\qquad \begin{cases} Y_{ww} = \Delta_1^{-1}(Y_{s,ww} + 2Y_{s,wx}\overline{X}_w + Y_{s,xx}\overline{X}_w^2 + Y_{s,x}A_{1,yy}Y_w^2), \\ Y_{wx} = \Delta_1^{-1}(Y_{s,wx}\overline{X}_x + Y_{s,xx}\overline{X}_w\overline{X}_x + Y_{s,x}Y_w(A_{1,xy} + A_{1,yy}Y_x)), \\ Y_{xx} = \Delta_1^{-1}(Y_{s,xx}\overline{X}_x^2 + Y_{s,x}(A_{1,xx} + 2A_{1,xy}Y_x + A_{1,yy}Y_x^2)), \\ Y_{wt} = \Delta_1^{-1}[Y_{s,wt} + Y_{s,wx}\overline{X}_t + Y_{s,xx}\overline{X}_w\overline{X}_t + Y_{s,xt}\overline{X}_w \\ \qquad + Y_{s,x}Y_w(A_{1,yt} + A_{1,yy}Y_t)], \\ Y_{xt} = \Delta_1^{-1}[Y_{s,xt}\overline{X}_x + Y_{s,xx}\overline{X}_x\overline{X}_t \\ \qquad + Y_{s,x}(A_{1,xt} + A_{1,xy}Y_t + A_{1,yt}Y_x + A_{1,yy}Y_xY_t)]. \end{cases}$$

A.3 *Estimates for simple composition.* — We keep the notations of Section A.1 above and use also the notations $|P|$, $|Q|$, $|P'|$, $|Q'|$ introduced in Section 3.2. Constants depending only on the cone condition satisfied by F_t, F'_t are denoted by C_0, those depending also on the distortions of F_t, F'_t by C.

From (A.6), we get

$$\text{(A.65)}\qquad \begin{cases} |X_x| \leqslant C_0|P'|, \\ |X_y| \leqslant C_0|Q|, \\ |Y_x| \leqslant C_0|P'|, \\ |Y_y| \leqslant C_0|Q|, \\ |X_t| \leqslant C_0(|A'_t| + |B_t|), \\ |Y_t| \leqslant C_0(|A'_t| + |B_t|). \end{cases}$$

Then (A.9) gives

$$\text{(A.66)}\qquad \begin{cases} |A''_y - A_y| \leqslant C_0|P||Q|, \\ |B''_x - B'_x| \leqslant C_0|P'||Q'|, \end{cases}$$

and (A.10) gives

$$\text{(A.67)}\qquad \begin{cases} |A''_t - A_t| \leqslant C_0|P|(|A'_t| + |B_t|), \\ |B''_t - B'_t| \leqslant C_0|Q'|(|A'_t| + |B_t|). \end{cases}$$

Next, (A.11) gives

$$\text{(A.68)}\qquad \begin{cases} |\Delta_x| \leqslant C_0|P'|(|B_{xx}| + |A'_{yy}| + |\partial_y \log|A'_x||) \leqslant C|P'|, \\ |\Delta_y| \leqslant C_0|Q|(|B_{xx}| + |A'_{yy}| + |\partial_x \log|B_y||) \leqslant C|Q|, \\ |\Delta_t| \leqslant C(|B_{xt}| + |A'_{yt}| + |A'_t| + |B_t|). \end{cases}$$

From (A.20)–(A.23) we obtain

$$\text{(A.69)}\qquad \begin{cases} |X_{yy}| \leqslant C_0|Q|(|A'_{yy}| + |B_{xx}| + |\partial_x \log|B_y|| + |\partial_y \log|B_y||) \leqslant C|Q|, \\ |Y_{xx}| \leqslant C_0|P'|(|B_{xx}| + |A'_{yy}| + |\partial_y \log|A'_x|| + |\partial_x \log|A'_x||) \leqslant C|P'|, \\ |X_{yt}| \leqslant C|Q|(|B_{xt}| + |A'_{yt}| + |A'_t| + |B_t| + |\partial_t \log|B_y||), \\ |Y_{xt}| \leqslant C|P'|(|B_{xt}| + |A'_{yt}| + |A'_t| + |B_t| + |\partial_t \log|A'_x||). \end{cases}$$

Then, from (A.12), (A.13), (A.15), (A.16), we get

$$\text{(A.70)}\qquad \begin{cases} |\partial_x \log|A''_x| - \partial_x \log|A'_x|| \leqslant C_0|P'|(D(F) + D(F')), \\ |\partial_y \log|A''_x| - \partial_y \log|A_x|| \leqslant C_0|Q|(D(F) + D(F')), \\ |\partial_x \log|B''_y| - \partial_x \log|B'_y|| \leqslant C_0|P'|(D(F) + D(F')), \\ |\partial_y \log|B''_y| - \partial_y \log|B_y|| \leqslant C_0|Q|(D(F) + D(F')), \end{cases}$$

while (A.18) gives

$$\text{(A.71)}\qquad \begin{cases} |A''_{yy} - A_{yy}| \leqslant C_0|P||Q|(D(F) + D(F')), \\ |B''_{xx} - B'_{xx}| \leqslant C_0|P'||Q'|(D(F) + D(F')). \end{cases}$$

Formulas (A.70) and (A.71) give the formula (3.13) for the distortion of a simple composition.

From (A.14), (A.17), we get

$$\text{(A.72)} \qquad \begin{cases} |\partial_t \log |A''_x| - \partial_t \log |A'_x| - \partial_t \log |A_x|| \leqslant C(|B_{xt}| + |A'_{yt}| + |A'_t| + |B_t|), \\ |\partial_t \log |B''_y| - \partial_t \log |B'_y| - \partial_t \log |B_y|| \leqslant C(|B_{xt}| + |A'_{yt}| + |A'_t| + |B_t|), \end{cases}$$

while (A.19) gives

$$\text{(A.73)} \qquad \begin{cases} |A''_{yt} - A_{yt}| \leqslant C|P|[|A'_t| + |B_t| + |Q|(|B_{xt}| + |A'_{yt}| + \partial_t \log |A_x| + \partial_t \log |B_y|)], \\ |B''_{xt} - B'_{xt}| \leqslant C|Q'|[|A'_t| + |B_t| + |P'|(|B_{xt}| + |A'_{yt}| + \partial_t \log |A'_x| + \partial_t \log |B'_y|)]. \end{cases}$$

Formulas (A.72), (A.73) are used in Section 7.4.

A.4 *Estimates for parabolic composition.* — The context and notations are those of Section A.2. We derive estimates from the formulas in this subsection, assuming that the maps F_0, F_1 satisfy (see (R4) in Section 5.3)

$$\text{(A.74)} \qquad \begin{cases} |A_{1,y}| < C\varepsilon_0, & |A_{1,yy}| < C\varepsilon_0, \\ |B_{0,x}| < C\varepsilon_0, & |B_{0,xx}| < C\varepsilon_0. \end{cases}$$

We write δ for $\delta(Q_0, P_1)$. We assume that (see (R7) in Section 5.4)

$$\text{(A.75)} \qquad \delta(Q_0, P_1) \geqslant C^{-1}\Big(|P_1|^{1-\eta} + |Q_0|^{1-\eta}\Big).$$

We first deal with the partial derivatives not involving time. From (A.29), we get

$$\text{(A.76)} \qquad \begin{cases} C^{-1} \leqslant |X_w| \leqslant C, \\ C^{-1} \leqslant |Y_w| \leqslant C, \\ |X_y| \leqslant C|Q_0|, \\ |Y_x| \leqslant C|P_1|. \end{cases}$$

Then, from (A.33), (A.34), we obtain

$$\text{(A.77)} \qquad \begin{cases} C^{-1}|P_1| \leqslant |\overline{X}_x| \leqslant C|P_1|, \\ C^{-1}|Q_0| \leqslant |\overline{Y}_y| \leqslant C|Q_0|, \\ |\overline{X}_w| \leqslant C\varepsilon_0, \\ |\overline{Y}_w| \leqslant C\varepsilon_0. \end{cases}$$

Putting this into (A.35) gives

$$\text{(A.78)} \qquad \begin{cases} C^{-1}|P_1| \leqslant |C_x| \leqslant C|P_1|, \\ C^{-1}|Q_0| \leqslant |C_y| \leqslant C|Q_0|. \end{cases}$$

From (A.63), (A.64), one gets

(A.79) $$\begin{cases} |X_{ww}| \leqslant C, \\ |Y_{ww}| \leqslant C, \end{cases}$$

and then, from (A.61), (A.62)

(A.80) $$\begin{cases} |\overline{X}_{ww}| \leqslant C\varepsilon_0, \\ |\overline{Y}_{ww}| \leqslant C\varepsilon_0. \end{cases}$$

Taking (A.77), (A.80) into (A.56) gives

(A.81) $$|C_{ww} - 2| \leqslant C\varepsilon_0.$$

Fix y_0, x_1 and denote by w^* the point where C takes its minimal value $\overline{C}(y_0, x_1)$. We have

(A.82) $$\begin{cases} |C_w(w, y_0, x_1) - 2(w - w^*)| \leqslant C\varepsilon_0|w - w^*|, \\ |C(w, y_0, x_1) - (w - w^*)^2 - \overline{C}(y_0, x_1)| \leqslant C\varepsilon_0|w - w^*|^2. \end{cases}$$

This gives, as $C(W(y_0, x_1), y_0, x_1) = 0$,

(A.83) $$||C_w(W(y_0, x_1), y_0, x_1)| - 2|\overline{C}(y_0, x_1)|^{1/2}| \leqslant C\varepsilon_0|\overline{C}(y_0, x_1)|^{1/2}.$$

From (A.75), (A.78) and the definition of $\delta = \min -C$, we have thus

(A.84) $$C^{-1}\delta^{1/2} \leqslant |C_w| \leqslant C\delta^{1/2}.$$

Formulas (3.27), (3.28) of Section 3.5 are then a consequence of (A.42).

From (A.41), we now get

(A.85) $$\begin{cases} C^{-1}|P_1|\delta^{-1/2} \leqslant |W_x| \leqslant C|P_1|\delta^{-1/2}, \\ C^{-1}|Q_0|\delta^{-1/2} \leqslant |W_y| \leqslant C|Q_0|\delta^{-1/2}. \end{cases}$$

Plugging into (A.39), (A.40) gives

(A.86) $$\begin{cases} |A_y - A_{0,y}| \leqslant C|P_0||Q_0|\delta^{-1/2}, \\ |B_x - B_{1,x}| \leqslant C|P_1||Q_1|\delta^{-1/2}. \end{cases}$$

Next, from (A.63), (A.64), we have

(A.87) $$\begin{cases} |X_{wy}| \leqslant C|Q_0|, \\ |X_{yy}| \leqslant C|Q_0|, \\ |Y_{wx}| \leqslant C|P_1|, \\ |Y_{xx}| \leqslant C|P_1|, \end{cases}$$

and then, from (A.61), (A.62)

$$\text{(A.88)} \qquad \begin{cases} |\overline{X}_{wx}| \leqslant C|P_1|, \\ |\overline{X}_{xx}| \leqslant C|P_1|, \\ |\overline{Y}_{wy}| \leqslant C|Q_0|, \\ |\overline{Y}_{yy}| \leqslant C|Q_0|. \end{cases}$$

Plugging into (A.57), (A.59) gives

$$\text{(A.89)} \qquad \begin{cases} |C_{wy}| \leqslant C|Q_0|, \\ |C_{yy}| \leqslant C|Q_0|, \\ |C_{wx}| \leqslant C|P_1|, \\ |C_{xx}| \leqslant C|P_1|, \\ |C_{xy}| \leqslant C|P_1||Q_0|. \end{cases}$$

We can now estimate the second partial derivatives of W from (A.55)

$$\text{(A.90)} \qquad \begin{cases} |W_{yy}| \leqslant C|Q_0|\delta^{-1/2}, \\ |W_{xx}| \leqslant C|P_1|\delta^{-1/2}, \\ |W_{xy}| \leqslant C|P_1||Q_0|\delta^{-1/2}, \end{cases}$$

which finally gives from (A.48), (A.53)

$$\text{(A.91)} \qquad \begin{cases} |A_{yy} - A_{0,yy}| \leqslant C|P_0||Q_0|\delta^{-1/2}, \\ |B_{xx} - B_{1,xx}| \leqslant C|P_1||Q_1|\delta^{-1/2}. \end{cases}$$

These estimates are used in Section 7.5. We now turn to derivatives involving the parameter. First, from (A.29), we get

$$\text{(A.92)} \qquad \begin{cases} |X_t| \leqslant C(1 + |B_{0,t}|), \\ |Y_t| \leqslant C(1 + |A_{1,t}|), \end{cases}$$

and then, from (A.33), (A.34)

$$\text{(A.93)} \qquad \begin{cases} |\overline{X}_t| \leqslant C(\varepsilon_0 + |B_{0,t}|), \\ |\overline{Y}_t| \leqslant C(\varepsilon_0 + |A_{1,t}|). \end{cases}$$

Plugging this into (A.35) gives

$$\text{(A.94)} \qquad |C_t| \leqslant C(1 + |B_{0,t}| + |A_{1,t}|),$$

and then, from (A.41)

$$|W_t| \leqslant C\delta^{-1/2}(1 + |B_{0,t}| + |A_{1,t}|). \quad \text{(A.95)}$$

From (A.44), one then concludes that

$$\begin{cases} |A_t - A_{0,t}| \leqslant C|P_0|\delta^{-1/2}(1 + |B_{0,t}| + |A_{1,t}|), \\ |B_t - B_{1,t}| \leqslant C|Q_1|\delta^{-1/2}(1 + |B_{0,t}| + |A_{1,t}|). \end{cases} \quad \text{(A.96)}$$

From now on, we assume that the estimates of Proposition 17 in Section 7.6 are satisfied, i.e.

$$\begin{cases} |A_{1,t}| \leqslant C\varepsilon_0^{1/2}, \\ |B_{0,t}| \leqslant C\varepsilon_0^{1/2}. \end{cases} \quad \text{(A.97)}$$

Observe then that in (A.92), (A.95), we get the estimates

$$\begin{cases} |X_t| \leqslant C, \\ |Y_t| \leqslant C, \\ |\overline{X}_t| \leqslant C\varepsilon_0^{1/2}, \\ |\overline{Y}_t| \leqslant C\varepsilon_0^{1/2}, \\ |C_t| \leqslant C, \\ |W_t| \leqslant C\delta^{-1/2}. \end{cases} \quad \text{(A.98)}$$

From (A.63), (A.64), we obtain

$$\begin{cases} |X_{wt}| \leqslant C(1 + |B_{0,xt}|), \\ |X_{yt}| \leqslant C|Q_0|(1 + |B_{0,xt}| + |\partial_t \log |B_{0,y}|), \\ |Y_{wt}| \leqslant C(1 + |A_{1,yt}|), \\ |Y_{xt}| \leqslant C|P_1|(1 + |A_{1,yt}| + |\partial_t \log |A_{1,x}|), \end{cases} \quad \text{(A.99)}$$

and then, from (A.61), (A.62)

$$\begin{cases} |\overline{Y}_{wt}| \leqslant C(\varepsilon_0 + |B_{0,xt}|), \\ |\overline{Y}_{yt}| \leqslant C|Q_0|(1 + |B_{0,xt}| + |\partial_t \log |B_{0,y}|), \\ |\overline{X}_{wt}| \leqslant C(\varepsilon_0 + |A_{1,yt}|), \\ |\overline{X}_{xt}| \leqslant C|P_1|(1 + |A_{1,yt}| + |\partial_t \log |A_{1,x}|). \end{cases} \quad \text{(A.100)}$$

Taking this into (A.58), (A.60) gives

$$\begin{cases} |C_{wt}| \leqslant C(\varepsilon_0 + |B_{0,xt}| + |A_{1,yt}|), \\ |C_{yt}| \leqslant C|Q_0|(1 + |B_{0,xt}| + |\partial_t \log |B_{0,y}|), \\ |C_{xt}| \leqslant C|P_1|(1 + |A_{1,yt}| + |\partial_t \log |A_{1,x}|). \end{cases} \quad \text{(A.101)}$$

We can now estimate the partial derivatives of W from (A.55)

$$\text{(A.102)}\quad \begin{cases} |W_{xt}| \leqslant C|P_1|\delta^{-1/2}[\delta^{-1} + \delta^{-1/2}(|B_{0,xt}| + |A_{1,yt}|) + |\partial_t \log|A_{1,x}||], \\ |W_{yt}| \leqslant C|Q_0|\delta^{-1/2}[\delta^{-1} + \delta^{-1/2}(|B_{0,xt}| + |A_{1,yt}|) + |\partial_t \log|B_{0,y}||]. \end{cases}$$

Finally, we obtain from (A.49), (A.54)

$$\text{(A.103)}\quad \begin{cases} |A_{yt} - A_{0,yt}| \leqslant C|P_0|\delta^{-1/2} + C|P_0||Q_0|\delta^{-1/2}[\delta^{-1} + \delta^{-1/2}(|B_{0,xt}| + |A_{1,yt}|) \\ \qquad\qquad + |\partial_t \log|B_{0,y}| + |\partial_t \log|A_{0,x}||], \\ |B_{xt} - B_{1,xt}| \leqslant C|Q_1|\delta^{-1/2} + C|P_1||Q_1|\delta^{-1/2}[\delta^{-1} + \delta^{-1/2}(|B_{0,xt}| + |A_{1,yt}|) \\ \qquad\qquad + |\partial_t \log|B_{1,y}| + |\partial_t \log|A_{1,x}||], \end{cases}$$

and from (A.47), (A.52)

$$\text{(A.104)}\quad \begin{cases} |\partial_t \log|A_x| - \partial_t \log|A_{0,x}|| \leqslant C[\delta^{-1} + \delta^{-1/2}(|B_{0,xt}| + |A_{1,yt}|) + |\partial_t \log|A_{1,x}||], \\ |\partial_t \log|B_y| - \partial_t \log|B_{1,y}|| \leqslant C[\delta^{-1} + \delta^{-1/2}(|B_{0,xt}| + |A_{1,yt}|) + |\partial_t \log|B_{0,y}||]. \end{cases}$$

These formulas are used in Section 7.7.

Appendix B: On the Lipschitz regularity of $\widetilde{R}_+^\infty$

The goal of this Appendix is to perform some calculations that are used in Section 10.5, Proposition 51.

We recall the setting We have rectangles $R_0 = I_0^s \times I_0^u$, $R_{a_s} = I_u^s \times I_u^u$, $R_{a_u} = I_s^s \times I_s^u$ with respective coordinates (x_0, y_0), (x_u, y_u), (x_s, y_s).

We have an affine-like map F with domain $P \subset R_0$, image $Q \subset Q_u \subset R_{a_u}$, and implicit representation (A, B). We assume as usual that the cone condition is satisfied and that the distortion is bounded by C. 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

$$\text{(B.1)}\quad \begin{cases} y_s = Y_s(w, x_s), \\ x_u = X_u(w, y_u), \\ w^2 = \theta(y_u, x_s) \end{cases}$$

with Y_s, X_u, θ as in Sections 2.3 and 3.5. One has (cf. (R4) in Section 5.3):

$$\text{(B.2)}\quad |B_x| < C\varepsilon_0, \qquad |B_{xx}| < C\varepsilon_0.$$

Consider a vertical-like C^2 curve $\omega = \{x_s = \varphi(y_s)\} \subset P_s \subset R_{a_s}$ intersecting L_s and satisfying

$$\text{(B.3)}\quad \left|\frac{\partial \varphi}{\partial y}\right| < C\varepsilon_0, \qquad \left|\frac{\partial^2 \varphi}{\partial y^2}\right| < C\varepsilon_0.$$

As in Section 3.5 (formulas (3.14), (3.15)) we eliminate y_u in the system $x_u = X_u(w, y_u)$, $y_u = B(y_0, x_u)$ to write

$$x_u = X(w, y_0). \tag{B.4}$$

Similarly, in the equation

$$y_s = Y_s(w, \varphi(y_s)), \tag{B.5}$$

we solve for y_s to write

$$y_s = Y(w). \tag{B.6}$$

We set

$$C(w, y_0) = w^2 - \theta(B(y_0, X(w, y_0)), \varphi(Y(w))). \tag{B.7}$$

We assume that, for all $y_0 \in I_0^u$

$$\overline{C}(y_0) := \min_w C(w, y_0) \leqslant -C^{-1}|Q|^{1-\eta}. \tag{B.8}$$

Suppose now that we have two vertical-like C^2 curves $\omega_i = \{x_s = \varphi_i(y_s)\}$, $i = 0, 1$ as above satisfying (B.3), (B.8). We also assume that $\varphi_0(y_s) \neq \varphi_1(y_s)$ for all y_s and that we have for some $T > 0$, everywhere on I_s^u

$$\left| \frac{\partial}{\partial y} \log |\varphi_1 - \varphi_0| \right| \leqslant T. \tag{B.9}$$

For $i = 0, 1$, let $\Omega_i^{\pm}$ be the connected components of $F^{-1}(Q \cap G^{-1}(\omega_i \cap L_s))$.

Proposition. — The curves $\Omega_i^{\pm}$ *are graphs*

$$\Omega_i^{\pm}(s) = \{x_0 = \Phi_i^{\pm}(y_0)\} \tag{B.10}$$

which satisfy, everywhere on I_0^u

$$\left| \frac{\partial}{\partial y} \log |\Phi_1^+ - \Phi_0^+| \right| \leqslant T', \tag{B.11}$$

$$\left| \frac{\partial}{\partial y} \log |\Phi_1^- - \Phi_0^-| \right| \leqslant T', \tag{B.12}$$

$$\left| \frac{\partial}{\partial y} \log |\Phi_1^+ - \Phi_0^-| \right| \leqslant T'', \tag{B.13}$$

$$\left| \frac{\partial}{\partial y} \log |\Phi_1^- - \Phi_0^+| \right| \leqslant T'', \tag{B.14}$$

with $T' = C(1 + |Q|^{\frac{1}{2}} T)$, $T'' = C$.

Remark. — In (B.13), we allow $\omega_0 = \omega_1$.

Proof. — Let C_i, $i = 0, 1$, be the function defined by (B.7) for ω_i. One solves $C_i = 0$ to get

$$w = W_i^{\pm}(y_0), \tag{B.15}$$

$$\Phi_i^{\pm}(y_0) = A(y_0, X(W_i^{\pm}(y_0), y_0)). \tag{B.16}$$

Let $\varepsilon, \varepsilon' \in \{+, -\}$, $W_0 = W_0^{\varepsilon}$, $W_1 = W_1^{\varepsilon'}$, $\Phi_0 = \Phi_0^{\varepsilon}$, $\Phi_1 = \Phi_1^{\varepsilon'}$. Define, for $s \in [0, 1]$

$$W_s = sW_1 + (1 - s)W_0. \tag{B.17}$$

One has

$$\begin{aligned}(\Phi_1 - \Phi_0)(y_0)\\ = (W_1 - W_0)(y_0) \int_0^1 A_x(y_0, X(W_s(y_0), y_0))X_w(W_s(y_0), y_0)\,ds.\end{aligned} \tag{B.18}$$

Set

$$a_s(y_0) = A_x(y_0, X(W_s(y_0), y_0))X_w(W_s(y_0), y_0). \tag{B.19}$$

From (A.76), we have, for all s, y_0

$$C^{-1}|P| \leqslant |a_s(y_0)| \leqslant C|P|. \tag{B.20}$$

One has

$$\begin{aligned}\frac{\partial}{\partial y}\log|a_s| = \partial_y \log|A_x| + \partial_x \log|A_x|\left(X_y + X_w \frac{\partial W_s}{\partial y}\right)\\ + X_w^{-1}\left(X_{wy} + X_{ww}\frac{\partial W_s}{\partial y}\right).\end{aligned} \tag{B.21}$$

From (A.76), (A.79), (A.87) we get

$$\left|\frac{\partial}{\partial y}\log|a_s|\right| \leqslant C\left(1 + \left|\frac{\partial W_s}{\partial y}\right|\right). \tag{B.22}$$

But we see as for (A.85) that we have, for $i = 0, 1$

$$\left|\frac{\partial W_i}{\partial y}\right| \leqslant C|Q|\delta_i^{-1/2}, \tag{B.23}$$

where $\delta_i := -\max_{y_0}\min_w C(w, y_0)$. We have assumed that

$$\delta_i \geqslant C^{-1}|Q|^{1-\eta}. \tag{B.24}$$

It then follows from (B.22) that the logarithmic derivative of $|a_s|$ is bounded by C. From (B.20), the same is true for the logarithmic derivative of $|\int_0^1 a_s ds|$. To prove the estimates of the proposition, in view of (B.18), it remains to see that the logarithmic derivative of $|W_1 - W_0|$ is bounded by $C(1 + |Q|^{\frac{1}{2}}T)$ or C (depending whether ε and ε' are equal or not).

1. We assume first that $\varepsilon \neq \varepsilon'$.

In this case, from (B.24) and the quadratic behavior of C_i (cf. (A.82)), we have

$$(\textbf{B.25}) \qquad |W_1 - W_0| \geqslant C^{-1}|Q|^{\frac{1}{2}(1-\eta)}.$$

Combining this with (B.23), (B.24) shows that the logarithmic derivative of $|W_1 - W_0|$ is bounded by C.

2. We assume now that $\varepsilon = \varepsilon'$.

We may assume that

$$(\textbf{B.26}) \qquad |W_1 - W_0| \ll |Q|^{\frac{1}{2}(1-\eta)},$$

otherwise we conclude as the first case. We interpolate between the two curves ω_0, ω_1 defining

$$(\textbf{B.27}) \qquad \varphi(y_s, s) = s\varphi_1(y_s) + (1-s)\varphi_0(y_s).$$

We have

$$(\textbf{B.28}) \qquad \left|\frac{\partial}{\partial y}\log\left|\frac{\partial\varphi}{\partial s}\right|\right| \leqslant T.$$

In the equation

$$(\textbf{B.29}) \qquad y_s = Y_s(w, \varphi(y_s, s)),$$

we solve for y_s to write

$$(\textbf{B.30}) \qquad y_s = Y(w, s).$$

We then set

$$(\textbf{B.31}) \qquad \overline{X}(w, s) := \varphi(Y(w, s), s),$$

$$(\textbf{B.32}) \qquad C(w, y_0, s) := w^2 - \theta(\overline{Y}(w, y_0), \overline{X}(w, s)).$$

Then, one solves $C = 0$ to get

$$(\textbf{B.33}) \qquad w = W^{\varepsilon}(y_0, s).$$

One has

$$\frac{\partial W^{\varepsilon}}{\partial s} = -\frac{\partial C}{\partial s} C_w^{-1} = C_w^{-1}\theta_x\left(\frac{\partial\varphi}{\partial s} + \frac{\partial\varphi}{\partial y}\frac{\partial Y}{\partial s}\right). \tag{B.34}$$

From (B.29), (B.30), one obtains

$$\frac{\partial Y}{\partial s} = Y_{s,x}\frac{\partial\varphi}{\partial s}\left(1 - Y_{s,x}\frac{\partial\varphi}{\partial y}\right)^{-1}, \tag{B.35}$$

where, here and below $Y_{s,x}$ stands for the partial derivative of Y_s with respect to x (and not the second partial derivative of Y!). Plugging this into (B.34), one obtains

$$\frac{\partial W^{\varepsilon}}{\partial s} = C_w^{-1}\theta_x\frac{\partial\varphi}{\partial s}\left(1 - Y_{s,x}\frac{\partial\varphi}{\partial y}\right)^{-1}. \tag{B.36}$$

We now take the logarithmic derivative of this product of four terms with relation to y_0. We obtain a sum of four terms:

$$Z_1 = -C_w^{-1}[C_{ww}W_y + C_{wy}], \tag{B.37}$$

$$Z_2 = \partial_y \log|\theta_x|(\overline{Y}_y + \overline{Y}_w W_y) + \partial_x \log|\theta_x|\overline{X}_w W_y, \tag{B.38}$$

$$Z_3 = \left(\frac{\partial}{\partial y}\log\left|\frac{\partial\varphi}{\partial s}\right|\right)Y_w W_y, \tag{B.39}$$

$$Z_4 = \left(1 - Y_{s,x}\frac{\partial\varphi}{\partial y}\right)^{-1}\left\{Y_{s,x}\frac{\partial^2\varphi}{\partial y^2}Y_w W_y + \frac{\partial\varphi}{\partial y}(Y_{s,xw}W_y + Y_{s,xx}\overline{X}_w W_y)\right\}. \tag{B.40}$$

The partial derivatives of X have been estimated in Appendix A (formulas (A.76), (A.79), (A.87))

$$|X_w| \leqslant C, \qquad |X_y| \leqslant C|Q|, \qquad |X_{ww}| \leqslant C, \qquad |X_{wy}| \leqslant C|Q|. \tag{B.41}$$

The partial derivatives of Y are estimated by

$$|Y_w| \leqslant C, \qquad |Y_{ww}| \leqslant C. \tag{B.42}$$

Then we get

$$|\overline{X}_w| \leqslant C\varepsilon_0, \qquad |\overline{Y}_w| \leqslant C\varepsilon_0, \qquad |\overline{Y}_y| \leqslant C|Q|, \tag{B.43}$$

$$|\overline{X}_{ww}| \leqslant C\varepsilon_0, \qquad |\overline{Y}_{ww}| \leqslant C\varepsilon_0, \qquad |\overline{Y}_{wy}| \leqslant C|Q|, \tag{B.44}$$

$$|C_{wy}| \leqslant C|Q|, \qquad |C_{ww} - 2| \leqslant C\varepsilon_0, \tag{B.45}$$

$$|W_y| \leqslant C|Q||C_w^{-1}|. \tag{B.46}$$

From (B.26), we have $|\delta_0 - \delta_1| \ll \delta_0$, and therefore we have

$$|C_w^{-1}(W^\varepsilon(y_0, s), y_0, s)| \leqslant C\delta_0^{-1/2}, \tag{B.47}$$

for all y_0, s. This gives, as $\delta_0 \geqslant C^{-1}|Q|^{1-\eta}$

$$|Z_i| \leqslant C \quad \text{for } i = 1, 2, 4; \tag{B.48}$$

$$|Z_3| \leqslant C|Q|^{1/2}T. \tag{B.49}$$

This proves that

$$\left|\frac{\partial}{\partial y}\log\left|\frac{\partial W^\varepsilon}{\partial s}\right|\right| \leqslant C(1 + |Q|^{1/2}T). \tag{B.50}$$

But then, we have

$$\begin{aligned}
\left|\frac{\partial}{\partial y}(W_1^\varepsilon - W_0^\varepsilon)\right| &= \left|\frac{\partial}{\partial y}\int_0^1 \frac{\partial W^\varepsilon}{\partial s} ds\right| \\
&\leqslant C(1 + |Q|^{1/2}T)\int_0^1 \left|\frac{\partial W^\varepsilon}{\partial s}\right| ds \\
&= C(1 + |Q|^{1/2}T)|W_1^\varepsilon - W_0^\varepsilon|,
\end{aligned}$$

as $\frac{\partial W^\varepsilon}{\partial s}$ has constant sign.

The required estimate on the logarithmic derivative of $|W_1 - W_0|$ has been obtained in both cases. The proof of the proposition is complete. □

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 Section 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 $\{x \geqslant x_0\}$ is finite and totally ordered. A *tree* is a forest with a single maximal element. Let $X_1, \ldots, X_n$ be forests and let A be a subset of $X = X_1 \times \cdots \times X_n$ (one should think of A as the graph of an *n-ary* relation). We say that A is *hereditary* if whenever $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ are such that and $y_i \leqslant x_i$ for all $1 \leqslant i \leqslant n$ (abbreviated as $y \leqslant x$), then $y \in A$ if $x \in A$.

Two points $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ of X are *coordinate-wise comparable* (*c-comparable* for short) if for each $i \in \{1, \ldots, n\}$, we have $x_i \geqslant y_i$ or $x_i \leqslant y_i$. In this case, we set

(**C.1**) $$x \vee y = (\max(x_i, y_i))_{1 \leqslant i \leqslant n}.$$

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 $x \leqslant y$ 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* $t_1 \leqslant x_1$, $t_2 \leqslant x_2$ *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. — 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 $y_1 \leqslant z_1$ and $y_2 \geqslant z_2$, and also $z_1 = z'_1, y_2 = y'_2$; then, we have $z'_1 = z_1 \geqslant y_1 \geqslant y'_1$ and $y'_2 = y_2 \geqslant z_2 \geqslant z'_2$, hence y', z' are c-comparable with $y' \vee z' = y \vee z$. □

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 $x'_i \geqslant t'_i$, $x''_i \geqslant t''_i$ 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. □

C.4 For $n \geqslant 3$, the situation is more complicated, as the two examples below indicate.

Example. — 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**) $$x_1 \geqslant y_1 \geqslant z_1,$$

(**C.3**) $y_2 \geqslant x_2, \qquad y_2 \geqslant z_2,$

(**C.4**) $z_3 \geqslant x_3, \qquad z_3 \geqslant y_3.$

Let $A = \{x, y, z\} \subset X$. If we define, as in the proposition above,

(**C.5**) $A_1 = \{u \vee v, u, v \in A, u, v\ c\text{-comparable}\}$

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**) $A_1 = \{x, y, z, y \vee z = (y_1, y_2, z_3)\}.$

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 $w_i \leqslant u_i$ $(i = 1, 2, 3)$ for any $u \in A_1$. This example shows that the analogue of the proposition above is false for $n = 3$.

Example. — 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**) $x_1 \geqslant y_1, \qquad x_1 \geqslant z_1,$

(**C.8**) $y_2 \geqslant x_2, \qquad y_2 \geqslant z_2,$

(**C.9**) $z_3 \geqslant x_3, \qquad z_3 \geqslant y_3,$

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 $\{u \leqslant x\}$, $\{v \leqslant y\}$, $\{w \leqslant z\}$ are disjoint and their union is the $c.h$-envelope of A: any $u \leqslant x$, $v \leqslant y$ 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 $u' \leqslant x$, $u'' \leqslant x$ and u', u'' c-comparable, then $u' \vee u'' \leqslant x$.

C.5 We have the following partial result:

Proposition. — *Let* $X_1, \ldots, X_n$ *be forests and let* A *be a subset of* $X = X_1 \times \cdots \times X_n$. *Let* A_1 *be the set of elements* $x \in X$ *for which there exists* $x^1, x^2, \ldots, x^n$ *in* A *with* $x_j = x_j^j \geqslant x_j^i$ *for all* $1 \leqslant i, j \leqslant n$. *Let* A_2 *be the set of elements* $y \in X$ *such that* $y \leqslant x$ *for some* x *in* A_1. *Then* A_2 *is hereditary and concave.*

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 Section C.3 above and left to the reader. □

C.6 We will now see how the definition of the transversality relation in Section 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 $\mathcal{R}(\mathrm{I})$. So, let us just try to define the relation for the starting class $\mathcal{R}(\mathrm{I}_0)$ associated to the initial horseshoe K. We would have:

$$X_1 = \{(P, Q, n) \in \mathcal{R}(I_0), Q \subset Q_u\}, \tag{C.10}$$

$$X_2 = \{(P', Q', n') \in \mathcal{R}(I_0), P \subset P_s\}, \tag{C.11}$$

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

$$Q \,\widehat{\pitchfork}_I\, P'$$

if for all $t \in I$ we have

$$\delta(Q, P') \geqslant 2 \max(I, |Q|^{1-\eta}, |P'|^{1-\eta}). \tag{C.12}$$

(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

$$A = \{(Q, P', I), Q \,\widehat{\pitchfork}_I\, P'\}. \tag{C.13}$$

This set is hereditary but it is not, a priori, concave. The concavity property (Proposition 9 in Section 6.3) 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 Section 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

$$Q \,\widehat{\pitchfork}_{I_1}\, P'_1, \qquad Q_2 \,\widehat{\pitchfork}_{I_2}\, P', \qquad Q_3 \,\widehat{\pitchfork}_I\, P'_3. \tag{C.14}$$

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

$$\delta(Q, P'_1) \geqslant 2 \max(|Q|^{1-\eta}, |P'_1|^{1-\eta}) \tag{C.15}$$

- there exists $Q_2 \subset Q$, $t_2 \in I$ such that

$$\delta(Q_2, P') \geqslant 2 \max(|Q_2|^{1-\eta}, |P'|^{1-\eta}) \tag{C.16}$$

– there exists $Q_3 \subset Q$, $P'_3 \subset P'$ such that

$$\delta(Q_3, P'_3) \geqslant 2 \max(|I|, |Q_3|^{1-\eta}, |P'_3|^{1-\eta}) \quad \textbf{(C.17)}$$

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

$$\delta(Q, P'_1) \geqslant 2|Q|^{1-\eta} \quad \text{for } t = t_1; \quad (C.15)'$$

$$\delta(Q_2, P') \geqslant 2|P'|^{1-\eta} \quad \text{for } t = t_2; \quad (C.16)'$$

$$\delta(Q_3, P'_3) \geqslant 2|I| \quad \text{for all } t \in I. \quad (C.17)'$$

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 Section 5.4. This defines $\overline{\pitchfork}$.

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 $\overline{\pitchfork}$ to $\pitchfork$ in Section 5.4.

REFERENCES

[BC] M. Benedicks and L. Carleson, The dynamics of the Hénon map, *Ann. Math.*, **133** (1991), 73–169.

[BDV] C. Bonatti, L. Diaz, and M. Viana, *Dynamics Beyond Uniform Hyperbolicity*, Encyclopedia of Math. Sciences, vol. 102, Springer, Berlin, 2004.

[BR] R. Bowen and D. Ruelle, The ergodic theory of Axiom A flows, *Invent. Math.*, **29** (1975), 181–202.

[C] E. Colli, Infinitely many coexisting strange attractors, *Ann. Inst. Henri Poincaré Anal. Non Linéaire*, **15** (1998), 539–579.

[CL] M. L. Cartwright and J. E. Littlewood, On nonlinear differential equations of the second order I, *J. Lond. Math. Soc.*, **29** (1945), 180–189.

[L] M. Levi, Qualitative analysis of the periodically forced relaxation oscillations, *Mem. Am. Math. Soc.*, **32** (1981), 244.

[MV] L. Mora and M. Viana, Abundance of strange attractors, *Acta Math.*, **171** (1993), 1–71.

[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** (2001), 475–480.

[MY] C. G. Moreira and J.-C. Yoccoz, Stable intersections of regular Cantor sets with large Hausdorff dimensions, *Ann. Math.*, **154** (2001), 45–96. Corrigendum in *Ann. Math.*, **154** (2001), 527.

[N] S. Newhouse, The abundance of wild hyperbolic sets and nonsmooth stable sets for diffeomorphisms, *Publ. Math. I.H.E.S.*, **50** (1979), 101–151.

[NP] S. Newhouse and J. Palis, Cycles and bifurcation theory, *Astérisque*, **31** (1976), 44–140.

[P1] J. Palis, A global view of Dynamics and a conjecture on the denseness of finitude of attractors, *Astérisque*, **261** (2000), 335–347. Géométrie complexe et systèmes dynamiques (Orsay, 1995).

[P2] J. Palis, A global perspective for non-conservative dynamics, *Ann. Inst. Henri Poincaré Anal. Non Linéaire*, **22** (2005), 485–507.

[P3] J. Palis, Open questions leading to a global perspective in dynamics, *Nonlinearity*, **21** (2008), 37–43.

[PT] J. Palis and F. Takens, Hyperbolic and the creation of homoclinic orbits, *Ann. Math.*, **125** (1987), 337–374.

[PY1] J. Palis and J.-C. Yoccoz, Homoclinic tangencies for hyperbolic sets of large Hausdorff dimension, *Acta Math.*, **172** (1994), 91–136.

[PY2] J. Palis and J.-C. Yoccoz, *Implicit Formalism for Affine-like Map and Parabolic Composition*, Global Analysis of Dynamical Systems, pp. 67–87, Institut of Phys., IOP, London, 2001.

[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** (2001), 867–871.

[Po] H. Poincaré, *Les méthodes nouvelles de la mécanique céleste*, vol. III, pp. 189, Gauthier-Villars, Paris, 1899.

[Ru] D. Ruelle, A measure associated with Axiom A attractors, *Am. J. Math.*, **98** (1976), 619–654.

[S] S. Smale, Differentiable dynamical systems, *Bull. Am. Math. Soc.*, **73** (1967), 747–817.

[Si] Ya. Sinai, Gibbs measures in ergodic theory, *Russ. Math. Surv.*, **27** (1972), 21–69.

J. P.
Instituto Nacional de Matemática Pura e Aplicada,
Estrada Dona Castorina, 110 Jardim Botânico,
22460-320, Rio de Janeiro RJ, Brazil
jpalis@impa.br

J.-C. Y.
Collège de France,
3 Rue d'Ulm,
75005 Paris, France
jean-c.yoccoz@college-de-france.fr

Manuscrit reçu le 27 avril 2006
publié en ligne le 15 octobre 2009.