TYPICAL LIMIT SETS OF CRITICAL POINTS FOR SMOOTH INTERVAL MAPS

Alexander Blokh and Michał Misiurewicz

August 22, 1997; revised June 10, 1998

Abstract. We prove that interval maps for which ω-limit sets of all critical points are minimal are dense in the space of all interval maps of class C^2.

1. Introduction

In this paper we continue our research on C^r-structurally stable interval maps for $r = 2$, started in [BM2] (see also [BM1]). It has been conjectured long ago that they satisfy Axiom A. For $r = 1$ this was proven by Jakobson [J]. Recently a proof in the unimodal case for all r has been announced by Kozlovski. However, the polymodal case is still open.

If the conjecture is true then for a dense set of maps the limit set (by this we mean the ω-limit set) of every critical point is finite. Here we prove a slightly weaker property, namely that for a dense set of maps the limit set of every critical point is minimal. In more technical terms, our main result can be stated as follows.

Main Theorem. *There is a dense subset $\mathcal{N}_2' \subset C^2([0,1],[0,1])$ such that if $f \in \mathcal{N}_2'$ then every critical point of f is either attracted to an attracting cycle or super persistently recurrent.*

Precise definitions will be given later. Here let us mention that super persistent recurrence is similar to persistent recurrence and implies minimality of the limit set (in the absence of wandering intervals). Thus, we get the following corollary to Main Theorem.

1991 *Mathematics Subject Classification.* 58F03, 58F10, 58F30.

Key words and phrases. Interval maps, structural stability.

The first author was partially supported by NSF grant DMS 9626303; the second author was partially supported by NSF grant DMS 9704422. We also would like to thank the referee for useful remarks.

Main Corollary. *There is a dense subset $\mathcal{N}_2' \subset C^2([0,1],[0,1])$ such that for any $f \in \mathcal{N}_2'$ the limit set of every critical point is minimal.*

The paper is organized as follows. In Section 2 we introduce the notion of super persistent recurrence and study its properties, using chains of intervals. In Section 3 we introduce several classes of maps and discuss what effects one can achieve with small perturbations of maps in these classes. In Section 4 we prove Main Theorem.

2. Topological properties of chains

One of the main tools we use in the paper are so-called *chains*. They were introduced by Lyubich in [L1] for interval maps with negative Schwarzian and helped him to prove non-existence of wandering intervals for such maps (an interval $J \subset [0,1]$ is called *wandering* for f if its images $f^n(J)$, $n \geq 0$, are pairwise disjoint and do not converge to a periodic orbit). Later chains were used to prove an analogous result for smooth polymodal interval maps (see [BL1], [MS, Chapter IV]; in [MS] they are called *pullbacks*) and became a popular tool in one-dimensional dynamics.

We assume that $f : [0,1] \to [0,1]$ is a piecewise monotone continuous map (strictly monotone on any lap). We call the local extrema of f (except 0 and 1) *turning points*. Let K_f be the closure of the convex hull of the union of the trajectories of the turning points of f. Clearly, K_f is a closed invariant interval. This is where the important things from the dynamical point of view happen. We want to have some extra space around K_f, so we assume that $0, 1 \notin K_f$. We call such f *loosely packed*. This assumption is not very restrictive, since in any reasonable topology loosely packed maps form an open dense subset of the appropriate space of interval maps. In fact, this assumption is not really necessary in order to make further definitions, but we make it for the sake of convenience.

Thus, throughout this section we assume that f is a piecewise monotone continuous loosely packed map of $[0,1]$ into itself. Moreover, we consider a finite set of points C containing $0, 1$ and all turning points of f. In the smooth case C is usually chosen as the set of all critical points of f with addition of 0 and 1 (in [BM2] such points are called *exceptional*). In the piecewise smooth case we usually also include in C all points at which the smoothness breaks down. However, we would like to emphasize that the constructions and results of this section hold for any set C containing $0, 1$ and all turning points of f, mainly because the definitions and arguments are topological.

We want to add here that some important concepts of the paper (e.g., that of a C-super persistently recurrent point) depend on the set C. That is, if $C \neq D$ then C-super persistent points of a map f may well be different from D-super persistently recurrent points. Therefore, although usually the set C comes with the map f and does not change, sometimes we consider various possibilities for this set with the fixed map f and choose among them the one that suits our purposes the best. Mimicking the terminology of [BM2] we call points from C *exceptional*.

If c is a turning point of f, let us take the largest interval $[a,b]$ such that $a < c < b$, $f(a) = f(b)$ and f is monotone on each of the intervals $[a,c]$ and $[c,b]$. Then there is a unique continuous function $\tau_c : [a, b] \to [a, b]$ such that $f \circ \tau = f$ and $f(x) \neq x$ for $x \neq c$. This function is an involution, that is τ_c^2 is the identity.

Now for given loosely packed f and C we choose a positive constant η such that

(A1) the distance between any two exceptional points of f is greater than η,
(A2) for any turning point c of f, the η-neighborhood of c is contained in the domain of τ_c,
(A3) for two exceptional points b, c either $f(b) = c$ or $|f(b) - c| > \eta$,
(A4) $K_f \subset (\eta, 1 - \eta)$.

Clearly, any sufficiently small η satisfies the above conditions.

Now we introduce the notion of a chain. The definitions we give are modifications of traditional definitions (see [L1], [MS, page 306]), designed to serve our purposes (for instance, we add (B3) below).

We call an interval *smart* if it does not contain any set of the form $f^k(V)$, where $k \geq 0$ and V is a one-sided η-neighborhood of an exceptional point of f. Note that any subinterval of a smart interval is also smart.

A sequence $(G_i)_{i=0}^l$ of closed intervals is called a *chain* if

(B1) G_i is a maximal interval such that $f(G_i) \subset G_{i+1}, i = 0, \ldots, l-1$,
(B2) $G_0 \cap K_f \neq \emptyset$,
(B3) G_l is smart.

The number l is called the *length* of the chain, G_0 is called the *first* interval of a chain, and G_l is called the *last* interval of the chain. The typical situation in which we deal with a chain of intervals is the following. Given a point x and an interval $I \ni f^n(x)$ we construct a chain of intervals $(G_i)_{i=0}^n$ whose last interval G_n is equal to I and whose first interval G_0 contains x. If such chain exists, it is unique. We call it the *pull-back chain of I along $x, \ldots, f^n(x)$* or just the *pull-back chain of I*. Any G_i is called a *pull-back* of I.

Construction of a pull-back is straightforward. Once we have G_i, we choose as G_{i-1} the component of $f^{-1}(G_i)$ containing $f^{i-1}(x)$. The only obstructions in the construction may be that (B2) or (B3) are not satisfied. However, condition (B2) is satisfied if $x \in K_f$. Condition (B3) says that I is smart. As we explain at the end of this section, in important cases this condition is satisfied.

When we have a chain $(G_i)_{i=0}^l$, we cannot avoid the situation when G_i contains exceptional points of f. However, we have the following lemma.

Lemma 2.1. *An interval G_i from a chain contains at most one exceptional point c of f, and if so then c is neither 0 nor 1. Moreover, if a turning point c of f belongs to G_i and $i < l$ then $\tau_c(G_i) = G_i$.*

Proof. By (B3), no interval G_i contains a one-sided η-neighborhood of an exceptional point of f. Thus, by (A1), G_i cannot contain more than one exceptional point of f. Suppose it contains one, and call it c. By (B2) and the invariance of K_f, we have $G_i \cap K_f \neq \emptyset$. Hence, by (A4), c cannot be 0 or 1. By (A2), G_i is contained in the domain of τ_c. If $i < l$ then from the maximality of G_i it follows that $\tau_c(G_i) = G_i$. ■

The intervals of a chain that contain elements of C play a special role. We call them C-*marked* (or just *marked*). Their number in a chain is called the C-*order* (or just *order*) of the chain. An exceptional point contained in a C-marked interval G_i is denoted by c_i. If c_i is a turning point of f then by Lemma 2.1 f is unimodal on G_i. Then the interval G_i is called also *unimodal*.

The next lemma follows immediately from Lemma 2.1 and the definition of a chain.

Lemma 2.2. *If G_i is a unimodal C-marked interval then $f|_{G_i}$ is unimodal, $f(G_i) \subset G_{i+1}$, and both endpoints of G_i are mapped by f into one endpoint of G_{i+1}. Otherwise, $f|_{G_i}$ is monotone, $f(G_i) = G_{i+1}$, and f maps the endpoints of G_i onto the endpoints of G_{i+1}.*

We will call an interval $I \subset [0,1]$ *nice* if for every $n > 0$ and an endpoint a of I the point $f^n(a)$ does not belong to the closure of I. In other words, the positive orbits of both endpoints of I miss the closure of I.

Lemma 2.3. *Let $(G_i)_{i=0}^l$ be the pull-back of G_l along $x, \dots, f^l(x)$ and let $0 < r \le l$. Assume that none of the points $x, \dots, f^{r-1}(x)$ belongs to G_r.*

(1) *Assume that G_l is nice. Then the intervals G_i, $0 \le i \le r$ are pairwise disjoint. In particular, C-order of the chain $(G_i)_{i=0}^r$ does not exceed the cardinality of C.*
(2) *Assume that the trajectories of the endpoints of G_l never enter the interior of G_l and that these endpoints are not postcritical. Then the interiors of the intervals G_i are pairwise disjoint. In particular, C-order of the chain $(G_i)_{i=0}^r$ does not exceed twice the cardinality of C.*

Proof. (1) Suppose that $a \in G_i \cap G_j$ for some $i < j \le r$. Since $f(G_k) \subset G_{k+1}$ for every k, we have $f^{r-j}(a) \in G_m \cap G_r$ for $m = r - j + i$. Since $i < j$, we have $m < r$. By the assumptions, $f^m(x)$ belongs to G_m, but not to G_r. Thus, one of the endpoints b of G_r belongs to G_m. By Lemma 2.2, $z = f^{l-r}(b)$ is an endpoint of G_l. The point $f^{r-m}(z) = f^{l-m}(b)$ belongs to G_l. This is a contradiction since G_l is nice and $r - m > 0$.

(2) In the proof of (1) the point b belongs to the interior of G_m. The only possibility for the point $f^{l-m}(b)$ to be an endpoint of G_l is that the trajectory of b passes through a critical point, but we assumed that the endpoints of G_l are not postcritical, so this is impossible. Thus, $f^{l-m}(b)$ belongs to the interior of G_l, and we get a contradiction as in the proof of (1). ∎

Now we are ready to introduce the notion of super persistent recurrence, which is similar to persistent recurrence (see e.g. [BL2], [BM2]). More precisely, we mean here the definition of persistent recurrence which relies upon the functions r_n (see below) and is different from Yoccoz' definition given for complex maps in terms of a puzzle. However, it is well-known (see, e.g., the discussion in [L2]) that for unimodal quadratic non-renormalizable maps the two definitions are equivalent.

We define super persistent recurrence by means of functions $r_n^k(x)$. The role they play is the same as the role of the functions r_n in the definition of persistent recurrence. Let us first recall the definition of r_n.

Let $f : [0,1] \to [0,1]$ be a piecewise monotone map. For $x \in [0,1]$ let us denote by $H_n(x)$ the maximal closed interval containing x on which f^n is monotone and let $f^n(H_n(x)) = M_n(x)$. Let $r_n(x)$ be the distance of $f^n(x)$ from the boundary of $M_n(x)$. If f^n has a local extremum at x, there is an ambiguity in the choice of $H_n(x)$ and $M_n(x)$, but $r_n(x) = 0$ independently of this choice. Moreover, in that case $r_m(x) = 0$ for all $m \geq n$. Also, if $x = 0$ or 1, then $r_n(x) = 0$ for all n. Thus either for some m we have $r_m(x) = 0$ (and then $r_n(x) = 0$ for all $n \geq m$) or $r_n(x) \neq 0$ for any n, in which case x is neither a preimage of a turning point nor $0, 1$. A recurrent point $x \in [0,1]$ is called *persistently recurrent* if $r_n(f(x)) \to 0$.

Now, let us fix the set C of exceptional points of f and consider the following construction. For every $\varepsilon > 0$ we construct the pull-back chain of $[f^n(x)-\varepsilon, f^n(x)+\varepsilon]$ along $x, \ldots, f^n(x)$ and denote by $m_{x,n}(\varepsilon)$ its order. Clearly, $m_{x,n}(\varepsilon)$ grows monotonically with ε. If there are no exceptional points among $x, f(x), \ldots, f^n(x)$ then for sufficiently small ε we have $m_{x,n}(\varepsilon) = 0$, otherwise even for arbitrarily small ε we have $m_{x,n}(\varepsilon) > 0$. We define $r_n^k(x)$ as the supremum of all ε such that $m_{x,n}(\varepsilon) \leq k$. Moreover, if for some N there are $k+1$ points of C among points $x, \ldots, f^N(x)$ then we define $r_n^k(x)$ as 0 for all $n \geq N$. In other words, positive $r_n^k(x)$ is the biggest number such that for every $\varepsilon' < r_n^k(x)$ the ε'-neighborhood of $f^n(x)$ can be pulled back along $x, \ldots, f^n(x)$ with order at most k. Note that $r_n^k(x)$ depends on f and C, yet for the sake of simplifying notation we avoid referring to them.

We call a recurrent point x such that for every k we have $r_n^k(x) \to 0$ *C-super persistently recurrent*. If we only claim the existence of a set C of exceptional points for which x is C-super persistently recurrent, but do not fix it, we call x simply a *super persistently recurrent* point.

Note that a point x is C-super persistently recurrent if and only if

$$\lim_{n\to\infty} m_{x,n}(\varepsilon) = \infty$$

for every $\varepsilon > 0$. This property is simpler than the one we introduced in the definition of C-super persistent recurrence. However, we defined C-super persistently recurrence by means of functions $r_n^k(x)$ because we wanted to stress the similarity of this notion to the usual persistent recurrence.

A set $A \subset [0,1]$ is called *minimal* if $f|_A$ is minimal.

In the following sections we shall prove that under the assumption of piecewise negative Schwarzian, the fact that a turning point of f is not C-super persistently recurrent for a specifically chosen set C implies that f is not stable. This will allow us to make conclusions about recurrent properties of critical points of piecewise negative Schwarzian maps, and eventually all interval maps of class C^2. However, in this section we are only concerned about topological properties of super persistently recurrent points. Namely, we prove the following theorem.

Theorem 2.4. *Let x be a super persistently recurrent point of f having arbitrarily small smart nice neighborhoods. Then $\omega(x)$ is minimal.*

Proof. If x is periodic then of course $\omega(x)$ is minimal. Let us assume that it is not periodic. Take C such that x is C-super persistently recurrent.

Suppose that $\omega(x)$ is not minimal. Then there exists a closed invariant non-empty set $A \subset \omega(x)$ not containing x. Hence, there exists a smart nice closed neighborhood $U = [a, b]$ of x, disjoint from A and shorter than η. The trajectory of x comes arbitrarily close to A, and therefore there are arbitrarily long pieces of this trajectory disjoint from U. Moreover, we can choose U so short that it does not contain points from C except maybe x (if $x \in C$) and that $f(x) \notin U$. Note that a and b do not belong to the trajectory of x since x is recurrent and U is nice.

For any $z \in U$ whose trajectory returns to U, denote by $\psi(z)$ the *first return time* into U. That is, $f^{\psi(z)}(z) \in U$, while $f^k(z) \notin U$ for $0 < k < \psi(z)$. Since x is recurrent, any point $f^n(x) \in U$ belongs to the domain of ψ. Since a and b do not belong to the trajectory of x, the function ψ is constant in a neighborhood of $f^n(x)$. We define *planks* as the maximal intervals containing points of the orbit of x on which ψ is constant (so if V is a plank it makes sense to talk about the first return time on V).

Let us establish some properties of planks. We claim that a plank is a closed interval. Indeed, if (u, v) is an interval on which ψ is equal to a constant l then by the continuity $f^l(u) \in U$ and on the other hand $f^k(u) \notin (a, b)$ for any $0 < k < l$. The former together with the niceness of U implies that also $f^k(u) \neq a, b$, so we see that $\psi(u) = l$ and similarly $\psi(v) = l$. This proves our claim. Thus, since ψ is not defined at the endpoints of U, every plank is contained in the interior of U.

If V is a plank with first return time l then by the maximality of V the endpoints of V are mapped by f^l to the endpoint(s) of U. Moreover, it is clear that different planks are disjoint.

Let us show that all planks are nice and smart. Let z be an endpoint of a plank V with first return time l. If $j < l$ then $f^j(z) \notin V$ by the definition of l. On the other hand as we have just seen $f^l(z) \in \{a, b\}$. Hence for $j \geq l$ also $f^j(z) \notin V$. Therefore V is nice. Smartness of V follows from smartness of U, since $V \subset U$.

Let us show also that a plank V with first return time l is a pull-back of U along $z, f(z), \dots, f^l(z)$ for any $z \in V$. Indeed, let $G_0, \dots, G_l = U$ be the corresponding pull-back chain. Then by Lemma 2.3 (1) intervals $G_1, \dots, G_l$ are pairwise disjoint, so $\psi|_{G_0} = l$ and $G_0 \subset V$. On the other hand, $f^l(V) \subset U$ and so $V \subset G_0$. Thus $G_0 = V$.

There are three types of chains related to planks. Let us fix a plank V. Let k be the smallest positive integer such that $f^k(x) \in V$. Then we can take the pull-back $(J, J', \dots, V)$ of V along $x, \dots, f^k(x)$ which we call a *chain of the first type*. By Lemma 2.3 (1), the order of the chain $(J', \dots, V)$ is not larger than the cardinality of C (denote this cardinality by N). Thus, the order of $(J, J', \dots, V)$ is at most $N + 1$.

Suppose that for a plank W there exists a point $c \in C$ such that the first time the trajectory of c enters U, it gets into W. In such a case we call W *exceptional* (we include also the possibility that $c \in W$). For every element of C we get at most one exceptional plank, so there are finitely many exceptional planks. Therefore there exists an interval I contained in the interior of U and containing all exceptional planks and a neighborhood of x.

Now let us take a plank V with first return time l and a point $f^k(x) \in V$. We will associate with this situation chains of the second and third type depending on the location of $f^{k+l}(x)$. Assume first that $f^{k+l}(x) \notin I$. The point $f^{k+l}(x)$ belongs to some plank W, so we can take the pull-back $(V', \dots, W)$ of W along $f^k(x), \dots, f^{k+l}(x)$ which we call a *chain of the second type*. If $f^{k+l}(x) \in I$, we pull back U (instead of W) along $f^k(x), \dots, f^{k+l}(x)$ to get a chain $(V'', \dots, U)$ which we call a *chain of the third type*. Since V is the first interval of the pull-back of U along $f^k(x), \dots, f^{k+l}(x)$ then in the case of the chain of the second type $V' \subset V$ and in the case of the chain of the third type $V'' = V$.

By the definition of I, the last interval of a chain of the second type is not exceptional. Therefore the order of a chain of the second type is 0. By Lemma 2.3 (1), the order of a chain of the third type is at most $N + 1$.

Now we construct a sequence of chains. We start with a plank V and the smallest k such that $f^k(x) \in V$; then we construct a chain of the first type $(J, \dots, V)$, which is a pull-back of a plank V along $x, \dots, f^k(x)$. Let l be the first return time on V. If $f^{k+l}(x) \in I$ then we take a chain of the third type $(V, \dots, U)$ and stop there. Otherwise, we take a chain of the second type $(V', \dots, W)$ with $V' \subset V$. Then we continue with W instead of V and $k+l$ instead of k (that is, we take a chain of the third or second type with the first interval contained in W, where the type of the chain we choose depends on whether the orbit of x hits or misses I when it returns to U next time). We continue like that until we take a chain of the third type. This has to happen, since x is recurrent, I contains a neighborhood of x, and thus the orbit of x enters I infinitely many times.

Let $f^m(x)$ be the last point of the orbit of x we deal with in the above construction. Then $f^m(x) \in I$. If we pull back U along $x, \dots, f^m(x)$ then we get a chain $\mathcal{G}$ with intervals contained in corresponding intervals of our sequence of chains. This is a kind of concatenation of chains. It is possible to produce it, since the last interval of a chain from the sequence contains the first interval of the next chain. The order of $\mathcal{G}$ is not larger than the sum of orders of the chains we concatenate. We used one chain of the first type, one of the third type and perhaps several chains of the second type. Therefore the order of $\mathcal{G}$ is at most $2N + 2$.

The number m above depends on the choice of the initial plank V. It is larger than the first return time on V. As we observed at the beginning of the proof, there are arbitrarily long pieces of the trajectory of x disjoint from U. Therefore there are planks with arbitrarily long first return time. Hence, m can be made arbitrarily large. Since U contains the ε-neighborhood of I, it contains the closed ε-neighborhood of $f^m(x)$. Therefore by reducing the size of the elements of $\mathcal{G}$, we can construct a pull-back of the closed ε-neighborhood of $f^m(x)$ along $x, \dots, f^m(x)$ of order at most $2N + 2$ for arbitrarily large m's. This contradicts the definition of C-super persistent recurrence of x. ■

We would like to use Theorem 2.4 to study the question of minimality of $\omega(c)$ in the case when c is a super persistently recurrent turning point of f. For this we need more definitions and lemmas.

A closed interval I is called *periodic (of period n)* if $I, \dots, f^{n-1}(I)$ are disjoint, while $f^n(I) \subset I$. Then the union $\bigcup_{i=0}^{n-1} f^i(I)$ is called a *cycle of intervals* and denoted by $\operatorname{cyc}(I)$. Clearly, if $J \subset I$ and both I, J are periodic then the period of J is a multiple of the period of I (yet these periods may well coincide). Let $I_0 \supset I_1 \supset \dots$ be a nested sequence of periodic intervals of periods $m_0 < m_1 < \dots$. Then the intersection $\bigcap_{i=0}^{\infty} \mathrm{cyc}(I_i)$ is called a *solenoidal set*. The dynamics of maps on solenoidal sets is well known. In particular, the following lemma holds.

Lemma 2.5 ([B]). *If x is contained in a solenoidal set then $\omega(x)$ is minimal.*

We need the following important fact. It follows easily from [B] but can be proven independently as well (see, e.g., [MS, page 305]).

Contraction Principle. *If I is an interval such that $\inf_n |f^n(I)| = 0$ then either I is a wandering interval or the orbit of I converges to a periodic orbit.*

If f has no wandering intervals (as for smooth maps, see [L1], [BL1], [MS, page 267]) then it is easy to find smart intervals.

Lemma 2.6. *Assume that f has no wandering intervals. Then every non-periodic point has a smart neighborhood.*

Proof. Suppose that x is not periodic, but has no smart neighborhood. Then there is a turning point c and its one-sided η-neighborhood V such that for every neighborhood U of x there is n with $f^n(V) \subset U$. Therefore $\inf_n |f^n(V)| = 0$. Since f has no wandering intervals, from Contraction Principle it follows that the sequence of intervals $(f^i(V))_{i=0}^{\infty}$ converges to a periodic orbit P. Since its subsequence converges to x, we get $x \in P$, a contradiction. This completes the proof. ■

Lemma 2.7. *Let c be a turning point of f. Assume that $\omega(c)$ is not a periodic orbit and that c does not belong to a solenoidal set. Then c has arbitrarily small nice neighborhoods.*

Proof. We want to show that for any neighborhood U of c there is a nice neighborhood of c contained in U. We may assume that U is small. Therefore we may assume that U does not contain any periodic interval containing c. Indeed, if there are arbitrarily small periodic intervals containing c then their periods either increase to infinity, and then c belongs to a solenoidal set, or stabilize, and then c is periodic.

Suppose that there is no nice neighborhood of c contained in U. If x is sufficiently close to c then $\tau_c(x)$ is defined and we can denote by I_x the closed interval with endpoints x and $\tau_c(x)$. By our assumption, the positive orbit of x enters I_x. Let m be the smallest positive integer such that $f^m(x) \in I_x$. Let us move x towards c. If m remains the same until x reaches c then $f^m(c) = c$, a contradiction. Therefore there is the last moment before m changes. For this value of x both endpoints of I_x are mapped to one of the endpoints by some $k \le m$. Thus, there is a periodic point $y \in I_x$ such that the trajectory of y never enters the interior of I_y. Since we could choose x arbitrarily close to c, we may assume that $I_y \subset U$.

If there is z in the interior of I_y and $n > 0$ such that $f^n(z)$ is y or $\tau_c(y)$ then we take the largest $i < n$ such that $f^i(z)$ is in the interior of I_y and then $I_{f^i(z)}$ is nice. This is a contradiction (we cannot have $f^i(z) = c$ since $\omega(c)$ is not a periodic orbit, so $I_{f^i(z)}$ is non-degenerate). Thus, if one point from the interior of I_y is mapped by

f^n to a point from I_y, so are all of them. Let us take $u \neq c$ from the interior of I_y. Since I_u is not nice, there is $n > 0$ such that $f^n(u) \in I_u \subset I_y$. Then $f^n(I_y) \subset I_y$, so I_y is a periodic interval, a contradiction. This completes the proof. ■

As an immediate consequence of Theorem 2.4 and Lemmas 2.5 - 2.7 we get the following theorem.

Theorem 2.8. *If f has no wandering intervals then any super persistently recurrent turning point of f has minimal limit set.*

Proof. We consider several cases. Let c be a super persistently recurrent turning point of f. If $\omega(c)$ is a periodic orbit then there is nothing to prove. If c belongs to a solenoidal set then by Lemma 2.5 the set $\omega(c)$ is minimal as well. Hence, we may assume that neither $\omega(c)$ is a cycle nor c belongs to a solenoidal set. Then by Lemma 2.7 c has arbitrary small nice neighborhoods. Since f has no wandering intervals we conclude by Lemma 2.6 that c has a smart neighborhood too. Therefore the conditions of Theorem 2.4 are satisfied for c and we conclude that $\omega(c)$ is minimal. ■

The next result of this section shows that super persistent recurrence of a turning point of a smooth map is basically a topological invariant. To state it we need the following definition. Consider the partition of $[0, 1]$ into turning points and the components of the complement of the set of turning points. The *itinerary* of $x \in [0, 1]$ for f is the sequence whose n-th term is equal to the element of the partition to which $f^n(x)$ belongs.

Theorem 2.9. *Let $g, h : [0, 1] \to [0, 1]$ be maps of class C^1 with finite number of turning points and no wandering intervals. Assume that they have the same turning points and for every turning point its itineraries for g and h coincide. Then every super persistently recurrent turning point of g with infinite limit set is super persistently recurrent for h and has infinite limit set for h.*

Proof. Since we are working with two maps at the same time, in order to distinguish between itineraries for g and h we will be speaking of *g-itineraries* and *h-itineraries* of points. From the kneading theory (see [MT]) we know that the sets of g-itineraries of all points of K_g and of h-itineraries of all points of K_h coincide, and the ordering of itineraries as we move from left to right is also the same for g and h. The sets of points with the same itinerary for a given map is connected. If such a set consists of one point, we call this point *thin*, otherwise such a set is an interval and we call its elements *thick*. Thus, the only difference that can occur between g and h is that there may be thin points for g that correspond to thick points for h or vice versa. However, even this is limited, since from the assumption on the absence of wandering intervals it follows that a thick point has a periodic orbit as the limit set. Moreover, a thick point is neither turning nor *preturning* (a preimage of a turning point under some iterate of the map).

Assume that C is a finite set containing all turning points of g. Assume also that a turning point c of g is C-super persistently recurrent for g with infinite limit set. Form a set D by taking for each element $x \in C$ one point with the same h-itinerary as the g-itinerary of x. Then D is a finite set containing all turning points of h. We are going to prove that c is D-persistently recurrent for h. We will denote by $m(f,n,\varepsilon)$ the order of the pull-back chain of $[f^n(c)-\varepsilon, f^n(c)+\varepsilon]$ along $c,\dots,f^n(c)$ (where $f=g,h$).

Suppose that c is not D-super persistently recurrent for h. Thus, there exists $\varepsilon>0$ and an integer k such that $m(h,n_i,\varepsilon)\le k$ for some increasing sequence (n_i) of integers. By passing to a subsequence, we may assume that both limits $x=\lim_{i\to\infty} h^{n_i}(c)$ and $y=\lim_{i\to\infty} g^{n_i}(c)$ exist.

For every n the g-itinerary of $g^n(c)$ is the same as the h-itinerary of $h^n(c)$. Therefore the g-itinerary of y and the h-itinerary of x coincide, unless for some $k\ge 0$ one of the points $h^k(x), g^k(y)$ is a turning point d while the other one belongs to an open interval of points with the same itineraries, whose one endpoint is d. By our assumption on the absence of wandering intervals, this means that the trajectories of x (for h) and y (for g) have finite limit sets. The point y belongs to the limit set A of c (for g). By Theorem 2.8, A is minimal, so the limit set of y is also A. By our assumptions A is infinite, a contradiction. This proves that the g-itinerary of y and the h-itinerary of x coincide.

Since the limit set of y for g infinite, y is thin. Moreover, the g-itinerary of y is neither periodic nor eventually periodic. The h-itinerary of x is the same, so x is thin for h.

If x is an endpoint of K_h then we replace each n_i by n_i+j for some j, which results in replacing y by $g^j(y)$ and x by $f^j(x)$. In the new situation our assumptions will be still satisfied, perhaps with smaller ε and larger k. On the other hand, with an appropriate choice of j the new point x will not be an endpoint of K_h (since none of the points $g^l(x)$ is periodic). Thus, we may assume that x is in the interior of K_h. We need that in order to be sure that the points in some neighborhood of x have their counterparts (in the sense of itineraries) in some neighborhood of y.

We claim that x is the limit from both sides of preturning points. Indeed, otherwise x is an endpoint of an interval of points with the same h-itineraries. This implies that the h-itinerary of x is periodic or eventually periodic, a contradiction. This proves our claim.

Hence, there are points a,b preturning for h, such that $a<x<b$ and the pull-back of $[a,b]$ along $c,\dots,h^{n_i}(c)$ has order at most k for every sufficiently large i. Let a' and b' be the points with the same g-itineraries as the h-itineraries of a and b respectively. Since a and b are preturning for h, the points a' and b' are uniquely determined and preturning for g. Moreover $a'<y<b'$. If i is sufficiently large then $a<h^{n_i}(c)<b$ and $a,h^{n_i}(c),b$ have distinct h-itineraries, so $a'<g^{n_i}(c)<b'$.

We claim that for such i the order of the pull-back of $[a',b']$ along $c,\dots,g^{n_i}(c)$ is the same as the order of the pull-back of $[a,b]$ along $c,\dots,h^{n_i}(c)$. If we compare corresponding intervals of both pull-back chains, the sets of the itineraries of their points are the same. Therefore the only differences could occur if an endpoint of such interval were thick with the same itinerary as some element of C. However, this is impossible, since by Lemma 2.2 the endpoints of all elements of both chains are preimages of the endpoints of $[a,b]$ or $[a',b']$. Therefore they are preturning, and consequently thin. This proves our claim.

There exists $\delta > 0$ such that for sufficiently large i the interval $[g^{n_i}(c)-\delta, g^{n_i}(c)+\delta]$ is contained in $[a', b']$. Therefore $m(g, n_i, \delta) \le k$, contrary to our assumption that c is C-persistently recurrent for g. Therefore c is D-super persistently recurrent for h.

To complete the proof, we have to show that c has infinite limit set for h. However, this follows from the fact that the limit set of a point is finite if and only if its itinerary is periodic or eventually periodic, and from our assumption that c has the same itinerary for g and h. ∎

We conclude this section with one more property of limit sets of super persistently recurrent points.

Proposition 2.10. *Let $X \subset [0,1]$ be an infinite minimal set for f. Assume that there are no wandering intervals. Then either every point of X is C-super persistently recurrent or no point of X is C-super persistently recurrent.*

Proof. Suppose that $x, y \in X$ and x is C-super persistently recurrent, while y is not. Then there exists $\varepsilon > 0$, a positive constant k and an increasing sequence (n_i) of positive integers such that $m_{y,n_i}(\varepsilon) \le k$ for every i. We may assume that the sequence $f^{n_i}(y)$ converges to some z. Then for every sufficiently large i there is a pull-back of $U' = [z-\varepsilon/2, z+\varepsilon/2]$ along $y, \dots, f^{n_i}(y)$ of order k or less.

If necessary, we replace U' by a smaller closed neighborhood of z which is disjoint from these closures of the trajectories of critical points that do not contain X. Since there are no wandering intervals, by Lemma 6.1 of [B] the set of points with finite limit sets is dense in the whole interval. Therefore we can find in U' one such point on each side of z. The union Y of the closures of the trajectories of these two points is disjoint from X. Let a, b be the points of Y closest to z from the left and right respectively. Then $z \in (a,b) \subset U'$ and the trajectories of a and b are disjoint from (a,b). Moreover, with our choice of U', the points a, b are not postcritical. Then the interval $U = [a,b]$ contains z in its interior and satisfies the conditions assumed for G_l in Lemma 2.3 (2).

Let us pull back U along $y, \dots, f^{n_i}(y)$ and denote by V_i the first interval of this pull-back. Let k_i be the smallest non-negative integer such that $f^{k_i}(x) \in V_i$ and let us prolong the above pull-back by pulling back V_i along $x, \dots, f^{k_i}(x)$. In such a way we get a chain that is a pull-back of U along $x, \dots, f^{k_i+n_i}(x)$, satisfying the assumptions of Lemma 2.3 (2) (with $G_r = V_i$ and $G_l = U$). Therefore $m_{x,k_i+n_i}(\delta)$, where δ is the minimal distance between points of X and Y, does not exceed twice the cardinality of C plus k. Since $\lim_{i\to\infty} n_i = \infty$, this contradicts the assumption that x is C-super persistently recurrent. ∎

Corollary 2.11. *If f has no wandering intervals and a turning point c of f is C-super persistently recurrent then every point of $\omega(c)$ is C-super persistently recurrent.*

3. Approximations and perturbations

In Sections 3 and 4 we will show that small C^2 perturbations of certain interval maps of class C^2 change their dynamics. Let us describe our approach step-by-step.

We begin with approximating our initial map f by a piecewise smooth map g with piecewise negative Schwarzian. Moreover, we make it in such a way that locally (in the sense of the appropriate topology in the space of piecewise-smooth maps) maximal possible number of critical points is attracted to attracting periodic orbits or mapped into other critical points (we call such maps *locally best*). It turns out that locally best maps have some additional useful properties, e.g., all their critical points which are neither attracted to attracting cycles or mapped into other critical points must be recurrent.

On the other hand, we show that there exists a C^2-smooth map h which is also close to f and has the same topological type as g; in other words, smoothness of g can be rebuilt without destroying its dynamics. Therefore if we can establish some additional topological properties of g then the same properties will be kept by h. To do so we rely upon Theorem 4.1 (which is the central result of Section 4), where we prove that if c is a critical point of a piecewise smooth map F with piecewise negative Schwarzian which is not C_F-super persistently recurrent (C_F is the set of all critical points of g together with all points of discontinuity of F'') then a small perturbation of F around c changes the itinerary of c. If $F = g$ is locally best, this is impossible. We conclude that all critical points of f (which are neither attracted to attracting cycles nor mapped into other critical points) are super persistently recurrent. The same behavior is exhibited by critical points of h, and h is a C^2-map from a small neighborhood of f. In such a way we get Main Theorem.

To make a preliminary step in which we approximate our initial map by a piecewise smooth one with piecewise negative Schwarzian and to be able to return afterwards to the space of maps of class C^2 we need approximation and perturbation lemmas, which thus become an important though rather technical tool for us. We prove them in this section. Some lemmas (specifically those devoted to studying so-called bump perturbations of piecewise negative Schwarzian maps) are simply borrowed from [BM2] where they have been proven for negative Schwarzian maps.

By $\mathcal{N}_r$ we denote the subspace of the space $C^r([0,1],[0,1])$ consisting of loosely packed maps with finitely many critical points, all of them non-degenerate, and none of them 0 or 1. It is well known (see e.g. [MS, page 217]) that the space of C^r maps with finitely many critical points, all of which are non-degenerate, is open and dense in $C^r([0,1],[0,1])$. The set of C^r maps for which 0 and 1 are not critical points is also open and dense. So is the set of loosely packed C^r maps, since if $f \in C^r([0,1],[0,1])$ then g given by $g(x) = (1-\varepsilon)f(x) + \varepsilon/2$ is loosely packed for every $\varepsilon > 0$. Hence, $\mathcal{N}_r$ is open and dense in $C^r([0,1],[0,1])$. In particular, $\mathcal{N}_2$ is a dense subset of $C^2([0,1],[0,1])$, which is in fact the first approximation result we will need.

Now, let us define maps with negative Schwarzian; we follow here [BM2]. Normally, one defines *Schwarzian* (or *Schwarzian derivative*) of a function f of class C^3 as $Sf = f'''/f' - (3/2)(f''/f')^2$. It is defined at all non-critical points of f. Thus, usually *negative Schwarzian* means $Sf < 0$ at all non-critical points. As can be easily checked, this property implies strict convexity of the function $1/\sqrt{|f'|}$ on each component of the complement of the set of critical points. This requires only C^1 smoothness, and we will adopt it as a definition of negative Schwarzian.

Thus, a function f is said to *have negative Schwarzian* if it is of class C^1 and the function $1/\sqrt{|f'|}$ is strictly convex on each component of the complement of the set of critical points. It is almost equivalent to the classical definition if f is C^3, and it is well known that it yields the same useful properties of interval maps.

Next we define two other similar spaces in which C^2-smoothness and negative Schwarzian are represented in a piecewise manner. The space $\mathcal{PN}$ consists of all maps $f : [0,1] \to [0,1]$ for which there exist points $0 = a_0 < a_1 < \cdots < a_{s-1} < a_s = 1$ such that

(1) f is of class C^1 on $[0,1]$,
(2) f has finitely many critical points,
(3) none of the points a_i is critical,
(4) f is loosely packed,
(5) f is of class C^2 on each (a_i, a_{i+1}) and can be extended to a function of class C^2 on $[a_i, a_{i+1}]$,
(6) all critical points of f are non-degenerate,

The space $\mathcal{PS}$ is defined in the same way with two additional conditions:

(7) f has negative Schwarzian on every interval $[a_i, a_{i+1}]$,
(8) f is of class C^3 in a neighborhood of all critical points.

These spaces are equipped with the C^2 topology. To define it precisely, we have to decide how to measure the distance between the second derivatives, since at some points only one-sided ones exist. To solve this problem, we set

$$\|f\|_2 = \max(\sup|f|, \sup|f'|, \sup|f_r''|, \sup|f_l''|),$$

where the supremum is taken over $[0,1]$ and f_r'' and f_l'' denote respectively the right and left one-sided second derivatives of f. Then the distance between f and g is $\|f-g\|_2$. Clearly, $\mathcal{N}_2$ and $\mathcal{PS}$ are subspaces of $\mathcal{PN}$.

A discontinuity of the second derivative of a function can be interpreted as a sudden change of acceleration. Therefore we will call $\sup_{x\in(0,1)} |f_r''(x) - f_l''(x)|$ the *kick* of f and denote it by $\operatorname{kick}(f)$. This notion is important if we want to approximate functions from $\mathcal{N}_2$ by functions from $\mathcal{PN}$ or vice versa.

The following property of approximations follows immediately from the definition of the C^2 distance and the triangle inequality.

Lemma 3.1. *If f and g are piecewise C^2 functions, then $|\operatorname{kick}(f) - \operatorname{kick}(g)| \leq 2\|f-g\|_2$. In particular, if $g \in \mathcal{N}_2$ then $\operatorname{kick}(g) = 0$, so $\operatorname{kick}(f) \leq 2\|f-g\|_2$.*

We want our approximations to preserve a part of dynamics. A point x is called *precritical* if $f^i(x)$ is critical for some $i > 0$ and *postcritical* if $x = f^j(c)$ for some critical point c and $j > 0$. We call a point x *p-critical* if it is either critical or both pre- and post-critical. More precisely, x is p-critical if there exists critical points c and d such that $x = f^i(c)$ and $f^j(x) = d$ for some $i, j \geq 0$. Clearly, the set of all p-critical points of a given map $f \in \mathcal{PN}$ is finite. We require that our approximations do not change the map at its p-critical points.

A point x is called a *periodic sink* (from one side) if there exists $n > 0$ and a (one-sided) neighborhood U of x such that $f^n(x) = x$, $f^n(U) \subset U$ and the diameter of $f^k(U)$ tends to 0 as $k \to \infty$. The *basin of attraction* of x is then the set $\bigcup_{k=0}^{\infty} f^{-k}(U)$. In this situation the number $(f^n)'(x)$, called *multiplier*, has absolute value less than or equal to 1 (recall that we are dealing with C^1-functions, therefore derivatives are defined in the usual, not just one-sided, sense). If the multiplier has the absolute value strictly less than 1 then x is called an *attracting periodic point* and its orbit is also called *attracting.* Finally, if the multiplier at a periodic point x has the absolute value 1 then the point x and its entire orbit are called *neutral.* A point x is *preperiodic* if $f^n(x)$ is periodic for some $n > 0$. We will refer to a point as *(pre)periodic* if it is either periodic or preperiodic. A point whose limit set is an attracting periodic orbit will be called *sinking.* If its limit set is a neutral periodic orbit, it will be called *weakly sinking.* If its limit set is a repelling periodic orbit which belongs to the boundary of the basin of attraction of a periodic sink, it will be called *almost sinking.* Otherwise, it will be called *floating.*

The next lemma is a kind of inverse of the second part of Lemma 3.1.

Lemma 3.2. *Let $f \in \mathcal{PN}$ and $\varepsilon > 0$ be such that* $\operatorname{kick}(f) < 2\varepsilon$. *Then there is $g \in \mathcal{N}_2$ such that $\|f - g\|_2 < \varepsilon$, $f(x) = g(x)$ for every p-critical point x of f, the maps f and g have the same critical points, and whenever a critical point is sinking for f, it is also sinking for g.*

Proof. We form a set A consisting of p-critical points of f and long pieces of trajectories of sinking critical points of f. We take these pieces so long that every point of discontinuity of f'' belonging to such trajectory is included, except those that are attracting periodic points themselves. Moreover, we follow such trajectory (and include its points in A) at least until we get to the basin of attraction of the attracting periodic orbit.

We have to modify f in small neighborhoods of the points of discontinuity of f'', to make it of class C^2. We have to preserve the value of the function at such a point, since this point may belong to A. If the neighborhood is sufficiently small, other elements of A lie outside it. Since f' does not vanish at the points of discontinuity of f'' (see Property (3) of maps from $\mathcal{PN}$), it has constant sign in such neighborhood $[a, b]$. We may assume that $f' > 0$ in $[a, b]$ and look for a function g of class C^2 on $[a, b]$ such that $\|g - f\|_2 < \varepsilon$, and g and f coincide at a and b up to their second derivatives. If c is the point of discontinuity of f'' in (a, b) then we require that $g(c) = f(c)$, $g'(c) = f'(c)$ and $g''(c) = (f''_r(c) + f''_l(c))/2$, and look separately at $[a, c]$ and $[c, b]$. The situation is similar on both intervals, so we consider only $[a, c]$.

Instead of approximating f by g, we will approximate f' by g'. The function g' has to have the same values at a and c as f'. Its derivative has to have the same value as the derivative of f' at a, but at c it has to be $\alpha = (f''_r(c) + f''_l(c))/2$. The assumption that $\operatorname{kick}(f) < 2\varepsilon$ implies that $|\alpha - f''_l(c)| < \varepsilon$. The function g' has to be ε-close to f' in C^1 and positive. This is clearly possible. Moreover, if we fix $\delta < \varepsilon$ such that $f' - \delta$ is still positive and δ is sufficiently small, we can find a version of g' that is equal to $f' - \delta$ on $((2a + c)/3, (a + 2c)/3)$ and is smaller than f' on an interval arbitrarily close to $[a, c]$ (to see that, look for $g' - f'$). This will give us a version of g' with the integral over $[a, c]$ smaller than the integral of f'. Similarly we find a version of g' with the integral larger than the integral of f'. By taking an appropriate convex combination we get a version of g' with the integral equal to the integral of f'. This will give us g that is equal to f at both a and c. The C^0 distance between g and f will be smaller than ε, since the same is true for g' and f' and the length of $[a, c]$ is smaller than 1. The rest of the conditions (except perhaps the one on the sinking critical points) are satisfied for both versions of g' and preserved by the convex combinations.

If a point of discontinuity of f'' is in the basin of attraction of an attracting periodic orbit, by taking sufficiently small neighborhood of this point for modification of the map, we can make the whole neighborhood to be contained in this basin of attraction (even after modification). Then the condition on the sinking critical points can be also satisfied. This completes the proof. ■

Our strategy includes approximating functions from $\mathcal{PN}$ by elements of $\mathcal{PS}$. The following lemma deals with this problem.

Lemma 3.3. *Let $f \in \mathcal{PN}$ and $\varepsilon > 0$ be given. Then there is $g \in \mathcal{PS}$ such that $\|f - g\|_2 < \varepsilon$, $f(x) = g(x)$ for every p-critical point x of f, the maps f and g have the same critical points, and whenever a critical point is sinking for f, it is also sinking for g.*

Proof. We start by finding $h \in \mathcal{PN}$ such that $\|f - h\|_2 < \varepsilon/2$, $f(x) = h(x)$ for every p-critical point x of f, the maps f and h have the same critical points, and h is of class C^3 in the neighborhoods of critical points. This requires modifications of the map in small neighborhoods of critical points. If such neighborhood is sufficiently small then the sign of f'' is constant on it and no p-critical point, except perhaps the critical point itself, belongs to it. Those modifications are so similar to the ones from the proof of the preceding lemma, that we leave them to the reader.

Now we have to approximate our map with one that has piecewise negative Schwarzian. Close to the critical points h has already negative Schwarzian (see e.g. [BM2]). Therefore we have to make modifications on some set Z where h' is bounded apart from 0. We may assume that Z is a union of finitely many intervals. Then we can divide Z into short intervals that do not contain p-critical points in their interiors. Let us take one of these intervals and call it $[a, b]$. Our aim is to find g on $[a, b]$ such that $\|h|_{[a,b]} - g\|_2 < \varepsilon/2$, $g(a) = h(a)$, $g(b) = h(b)$, $g'(a) = h'(a)$ and $g'(b) = h'(b)$, the function g is piecewise C^2 with negative Schwarzian on each piece and no critical points. If we can do it on each of our intervals, we get the required function g. We do not have to worry about the values of g'' at a and b, since these points may become points of discontinuity of g''. We may also assume that $[a, b]$ is as short as necessary.

Instead of working with h and g, we will work with $H = 1/\sqrt{|(h|_{[a,b]})'|}$ and $G = 1/\sqrt{|g'|}$. That is, H is given and we are looking for G. When we find it then we define g' as plus or minus the reciprocal of the square of G (the sign has to coincide with the sign of h'). Then we get g by integrating g'. In order to satisfy $g(a) = h(a)$ and $g(b) = h(b)$, we have to have equality of the integrals of h' and g' from a to b. This translates to the equality of the integrals of G^{-2} and H^{-2} (here superscripts mean power, not iterate). Next requirements mean that G is piecewise C^1 and strictly convex on each piece, and that $G(a) = H(a)$ and $G(b) = H(b)$. The inequality $\|h|_{[a,b]} - g\|_2 < \varepsilon/2$ will be satisfied if the C^1 distance between G and H is smaller than a certain positive constant δ. This δ depends only on ε and the bounds for the moduli of the derivatives of h on Z (the upper and lower bounds for $|h'|$ and the upper bound for $|h''|$).

Now we start to specify how fine the partition of Z has to be. Namely, H' has to vary less than by $\delta/2$ on each element of this partition. Recall that $[a,b]$ was an arbitrary element of this partition. We construct two continuous piecewise linear functions on $[a,b]$. Let us call them K_u and K_l. They both attain values $H(a)$ at a and $H(b)$ at b. They are linear (more precisely, affine) on $[a,d]$ and $[d,b]$, where d is the midpoint of $[a,b]$. Let $\alpha = (H(b) - H(a))/(b-a)$. Then K_u has slope $\alpha + \delta/2$ on $[a,d]$ and $\alpha - \delta/2$ on $[d,b]$, while K_l has slope $\alpha - \delta/2$ on $[a,d]$ and $\alpha + \delta/2$ on $[d,b]$. If our partition was sufficiently fine then both K_u and K_l are positive. Clearly, $K_l < H < K_u$ on (a,b), so $\int_a^b K_l^{-2}(x)\,dx > \int_a^b H^{-2}(x)\,dx > \int_a^b K_u^{-2}(x)\,dx$. Moreover, the C^1 distances between K_u and H and between K_l and H are smaller than δ. Now if we set $G_u(x) = K_u(x) + \beta(x-a)(x-b)$ and $G_l(x) = K_l(x) + \beta(x-a)(x-b)$ for sufficiently small β then these functions satisfy all conditions required for G, except the one about the integral. Moreover, $\int_a^b G_l^{-2}(x)\,dx > \int_a^b H^{-2}(x)\,dx > \int_a^b G_u^{-2}(x)\,dx$. Now we take as G the convex combination of G_u and G_l for which $\int_a^b G^{-2}(x)\,dx = \int_a^b H^{-2}(x)\,dx$. Clearly, it satisfies the rest of required conditions, except perhaps the one about the sinking critical points. However, this condition is satisfied automatically if the C^1 distance between f and g is sufficiently small. This completes the proof. ∎

Note that in the proof of Lemma 3.2 we could not use the same argument about the sinking critical points as at the end of the proof of Lemma 3.3, since the distance between f and g in Lemma 3.2 depends on $\mathrm{kick}(f)$.

In Section 4 our aim will be to make a small perturbation of a map in order to make the behavior of the trajectories of the critical points "better". Making a critical point precritical decreases the number of points about which we have to worry. Making a critical point sinking is the ultimate goal (perhaps not always achievable). Thus, we will use the following terminology. If $f, g \in \mathcal{PN}$ and either f has more precritical critical points than g or f has the same number of precritical critical points as g, but more sinking critical points than g, then we will say that f is *better* than g (and, of course, g is *worse* than f). A map $f \in \mathcal{PN}$ will be called *locally best* if it has a neighborhood consisting of maps that are not better than f. If $X \subset \mathcal{PN}$ and $f \in X$ is not worse than any element of X then we say that f is *best in* X. We do not say here *the* best, since usually there are many equally good maps.

It turns out that locally best maps have a kind of stability property.

Lemma 3.4. *Let f be best in a convex set $X \subset \mathcal{PN}$. If $g \in X$, $g(x) = f(x)$ for every p-critical point x of f, and the maps f and g have the same critical points, then for every critical point its itineraries for f and g coincide.*

Proof. Assume that $g \in X$, $g(x) = f(x)$ for every p-critical point x of f, and the maps f and g have the same critical points. Then every precritical critical point of f is precritical critical for g. This conclusion remains true if we replace g by $h = tf + (1-t)g$ for some $t \in [0,1]$. Suppose that a critical point c has different itineraries for f and g. We may assume that c is not precritical (otherwise we replace it by the last point on the orbit of c that is critical). Then for some t and h as above, the point c is precritical, and thus h is better than f, a contradiction. This completes the proof. ■

Now we prove two results that can be viewed as stronger versions of Lemmas 3.2 and 3.3. The first of them (Lemma 3.5) follows from Lemmas 3.1, 3.2 and 3.4.

Lemma 3.5. *Let X be the open ball in $\mathcal{PN}$ of radius $2\varepsilon > 0$ and center $f \in \mathcal{N}_2$. Assume that $g \in \mathcal{PN}$ is best in X and $\|f-g\|_2 < \varepsilon$. Then there is $h \in X \cap \mathcal{N}_2$ such that $h(x) = g(x)$ for every p-critical point of g, the maps g and h have the same critical points, the itineraries of critical points for g and h coincide, and whenever a critical point is sinking for g, it is also sinking for h (so that h is best in X as well).*

Proof. By Lemma 3.1 the fact that $\|f-g\|_2 < \varepsilon$ and $f \in \mathcal{N}_2$ implies that $\operatorname{kick}(g) < 2\varepsilon$. Then by Lemma 3.2 there is $h \in \mathcal{N}_2$ such that $\|h-g\|_2 < \varepsilon$, $h(x) = g(x)$ for every p-critical point of g, the maps h and g have the same critical points, and whenever a critical point is sinking for g, it is also sinking for h. Since g is best in X, we can apply Lemma 3.4, according to which for every critical point its itineraries for h and g coincide. This completes the proof. ■

It turns out that a map g from the preceding lemma can be picked up from the space $\mathcal{PS}$.

Lemma 3.6. *For every $f \in \mathcal{N}_2$ if $\varepsilon > 0$ is sufficiently small then there exists $g \in \mathcal{PS}$ satisfying the assumptions of Lemma 3.5.*

Proof. All critical points of f are nondegenerate, so if ε is sufficiently small then the number of critical points of all elements of the open ε-ball B_ε in $\mathcal{PN}$ centered at f is the same. Call this number N. Let $\xi(\widetilde{f})$ be the number of precritical critical points of $\widetilde{f}$ times N plus the number of sinking critical points of $\widetilde{f}$. As $\widetilde{f} \in X$, we have $\xi(\widetilde{f}) \le N^2 + N$. Thus $m = \limsup_{\widetilde{f} \to f} \xi(\widetilde{f})$ is finite. Therefore, as $\varepsilon \to 0$, the maximum of ξ over B_ε stabilizes. Hence, if ε is sufficiently small, this maximum is the same for ε and 2ε. We take such ε and a map $\widetilde{f} \in B_\varepsilon$ such that $\xi(\widetilde{f}) = m$. It is easy to see then that $\widetilde{f}$ is best in $X_{2\varepsilon}$.

Now by Lemma 3.3 we can approximate this $\widetilde{f}$ arbitrarily well by $g \in \mathcal{PS}$ with the same value of ξ. Therefore g is also best in X. If g is sufficiently close to $\widetilde{f}$ then $\|f-g\|_2 < \varepsilon$. Thus g satisfies the assumptions of Lemma 3.5. ■

Next, we need some tools from [BM2] that will allow us to make perturbations in neighborhoods of the critical points of a map. The proofs of Lemmas 3.7 and 3.9 can be found in [BM2], but we include them here to make the paper more self-contained.

Lemma 3.7 [BM2]**.** *For any $\delta, \varepsilon > 0$ there exists an even function $s_{\varepsilon,\delta} = s : \mathbb{R} \to \mathbb{R}$ of class C^3 such that*

(1) *$s(x) = 0$ for any $x \notin [-\varepsilon, \varepsilon]$, while $s(0) = \varepsilon^2\delta/1000$,*

(2) *s is strictly increasing on $[-\varepsilon, 0]$ and strictly decreasing on $[0, \varepsilon]$,*
(3) *$|s'(x)| < \varepsilon\delta$, $|s''(x)| \le \delta$ and $|s'''(x)| \le \delta/\varepsilon$ for any x.*

If $\varepsilon < 1$ then the C^2-norm of s is smaller than or equal to δ.

Proof. Set $h(x) = (x^2 - 1)^4$ for $x \in [-1, 1]$ and $h(x) = 0$ otherwise. This function satisfies the conditions of the lemma with $\varepsilon = 1$ and $\delta = 1000$. Now the function $s(x) = \delta\varepsilon^2 h(x/\varepsilon)/1000$ has all the required properties. ∎

In the typical situation both ε and δ are small, which guarantees that s is small in C^2 topology (yet to see whether s is small in C^3 topology, we need to know also how δ and ε are related). Since our main focus is C^2 topology, this allows us to perturb our maps by adding or subtracting functions similar to s.

Let $c \in \mathbb{R}$; we call the function $s_c(x) = s_{\varepsilon,\delta}(x - c)$ the *(ε, δ)-bump function at c*. If both ε and δ are small then we say that the (ε, δ)-bump function is small too. If c is a critical point of f, we can consider a map $g = f + s_c$ or $g = f - s_c$ where s_c is an (ε, δ)-bump function at c. Moreover, if ε is smaller than half the minimal distance between critical points of f then intervals supporting functions s_c for different critical points c are pairwise disjoint. From now on we consider functions s_c only for ε smaller than half the minimal distance between critical points of f. An *(ε, δ)-bump perturbation* of f is the result of adding to or subtracting from f some of maps s_c corresponding to different critical points c (the maps $f + s_c$ and $f - s_c$ are called *(ε, δ)-bump perturbations at c*). Note that by choosing small ε and δ we can get (ε, δ)-bump perturbations of f arbitrarily close to f in C^2 topology.

Observe that we can vary δ and then our bump perturbations depend continuously on it. Observe also that if no critical point of f is mapped into 0 or 1 (in particular, if f is loosely packed) then all sufficiently small bump perturbations at any critical point map $[0, 1]$ to itself.

The situation in which we apply bump perturbations will often be of the following type. For a non-degenerate critical point c of $f \in \mathcal{PN}$ we find points a close to c and b close to $f(c)$. Then we want to add to or subtract from f a bump function supported by a subset of $[a, \tau_c(a)]$ to get a new map g for which $g(c) = b$. The following lemma shows that under appropriate assumptions such perturbation may be made very small.

Lemma 3.8. *Let c be a critical point of a map $f \in \mathcal{PN}$. Then for any $\eta > 0$ there exist $\varepsilon, \delta > 0$ such that if $|a - \tau_c(a)| < \varepsilon$ and $0 < |b - f(c)|/|f(c) - f(a)| < \delta$ then there exists an $(\varepsilon_0, \delta_0)$-bump perturbation g of f such that $\varepsilon_0, \delta_0 < \eta$, the maps f and g coincide outside $[a; \tau_c(a)]$, and $g(c) = b$.*

Proof. Let $\gamma = \sup_{x\in[0,1]} |f''(x)|$. Set $\varepsilon = \eta$, $\delta = \eta/(500\gamma)$, $\alpha = \min(|a - c|, |\tau_c(a) - c|)$, $\varepsilon_0 = \alpha$ and $\delta_0 = 1000|b - f(c)|/\alpha^2$. By Lemma 3.7 the maximal value of the $(\varepsilon_0, \delta_0)$-bump function at c (call it s_c) is attained at c and equals $\varepsilon_0^2\delta_0/1000 = |b - f(c)|$. Set $g = f + s_c$ if $b > f(c)$ and $g = f - s_c$ if $b < f(c)$. Then $g(c) = b$. The support of $g - f$ is $[c - \varepsilon_0, c + \varepsilon_0] \subset [a; \tau_c(a)]$.

Assume that $|a - \tau_c(a)| < \varepsilon$ and $0 < |b - f(c)|/|f(c) - f(a)| < \delta$. We have to show that in this case $\varepsilon_0, \delta_0 < \eta$. We have $\varepsilon_0 = \alpha < |a - \tau_c(a)| < \varepsilon = \eta$. Furthermore, $f'(c) = 0$ implies that $|f'(x)| \le \gamma|x - c|$ for every x, and thus

$|f(a)-f(c)| \le \gamma|y-c|^2/2$ for $y = a, \tau_c(a)$. Therefore $|f(a)-f(c)| \le \gamma\alpha^2/2$. Since $|b-f(c)|/|f(c)-f(a)| < \delta = \eta/(500\gamma)$, we get $|b-f(c)| < \eta\alpha^2/1000$, so $\delta_0 < \eta$. This completes the proof. ■

Any bump perturbation has the same set of critical points as the original map, provided that ε and δ are sufficiently small and all the critical points of the original map are non-degenerate. Moreover, in the circumstances we will be using it, it preserves negative Schwarzian in neighborhoods of critical points, as the next lemma shows. This explains why we needed extra smoothness close to critical points in the definition of $\mathcal{PS}$.

We will say that f has *strongly negative Schwarzian* on an interval J if it is piecewise C^3 on J and there is $\varepsilon > 0$ such that $2f'''f' - 3(f'')^2 < -\varepsilon$ on J.

Lemma 3.9 [BM2]. *Let $f : [-a,a] \to \mathbb{R}$ be a function of class C^3 with a non-degenerate critical point 0. Then there exist positive numbers ε and δ such that for any C^3 function $s : [-a,a] \to \mathbb{R}$ whose support is contained in $[-\varepsilon,\varepsilon]$ and such that*

$$|s(x)| \le \delta, \quad |s'(x)| \le \varepsilon\delta, \quad |s''(x)| \le \delta, \quad |s'''(x)| \le \frac{\delta}{\varepsilon} \tag{3.7}$$

for all x, the function $g = f + s$ has strongly negative Schwarzian and a unique critical point in $[-\varepsilon,\varepsilon]$. In particular, this applies to any (ε,δ)-bump perturbation of f at 0. In this case, the critical point in $[-\varepsilon,\varepsilon]$ is 0.

Proof. Set

$$b = |f''(0)|, \quad d = \max_{x\in[0,1]} |f'''(x)|. \tag{3.8}$$

Since 0 is a non-degenerate critical point, we have $b > 0$. Choose $\varepsilon \in (0,a)$ such that

$$|f'(x)| < \min\left(1.3\, b|x|, 0.2\,\frac{b^2}{d}\right), \quad |f''(x)| > 0.8\, b \tag{3.9}$$

for $x \in [-\varepsilon,\varepsilon]$. Then choose $\delta > 0$ such that

$$\delta < \min\left(0.1\, b, 0.2\,\frac{b^2}{d\varepsilon}\right) \tag{3.10}$$

for $x \in [-\varepsilon,\varepsilon]$.

Let us check that these numbers have the required property. Let s be a function with the properties from the statement of the lemma and consider the function $g = f+s$. For $x \in [-\varepsilon,\varepsilon]$ we have $|g''(x)| \ge |f''(x)| - |s''(x)| \ge 0.8b - 0.1b \ne 0$, so g has at most one critical point in $[-\varepsilon,\varepsilon]$. It has to have one, since $g'(-\varepsilon) = f'(-\varepsilon)$ and $g'(\varepsilon) = f'(\varepsilon)$ have opposite signs. If s is a bump function, it is even, so $s'(0) = 0$. Hence, $g'(0) = f'(0) + s'(0) = 0$ and 0 is the only critical point of $g|_{[-\varepsilon,\varepsilon]}$.

Now we will check whether $g|_{[-\varepsilon,\varepsilon]}$ has strongly negative Schwarzian. To this end we show that

$$2(f'''(x)+s'''(x))(f'(x)+s'(x)) < 3(f''(x)+s''(x))^2 - 0.39\,b^2 \tag{3.11}$$

for $x \in [-\varepsilon, \varepsilon]$. In order to do this let us estimate both parts of the inequality (3.11) step by step. We will use all the time inequalities (3.7)-(3.10). We get

$$3(f''(x)+s''(x))^2 > 3(0.8\,b - 0.1\,b)^2 = 1.47\,b^2, \tag{3.12}$$

$$|f'''(x)| \cdot |f'(x)+s'(x)| < d\left(0.2\,\frac{b^2}{d} + \varepsilon \cdot 0.2\,\frac{b^2}{d\varepsilon}\right) = 0.4\,b^2, \tag{3.13}$$

$$|s'''(x)| \cdot |f'(x)+s'(x)| < \frac{0.1\,b}{\varepsilon} \cdot (1.3\,b\varepsilon + \varepsilon \cdot 0.1\,b) = 0.14\,b^2. \tag{3.14}$$

From (3.12)-(3.14) we get

$$\begin{aligned} 2(f'''(x)+s'''(x))(f'(x)+s'(x)) &< 2(0.4\,b^2 + 0.14\,b^2) \\ &= 1.08\,b^2 = 1.47\,b^2 - 0.39\,b^2 < 3(f''(x)+s''(x))^2 - 0.39\,b^2 \end{aligned}$$

for $x \in [-\varepsilon, \varepsilon]$. This proves (3.11) and completes the proof. ■

Note that due to Lemma 3.9, a map which is of class C^3 in a neighborhood of a non-degenerate critical point has strongly negative Schwarzian in a sufficiently small neighborhood of this point.

Now we return to the properties of locally best maps. To begin with we need two preliminary lemmas. The first one shows that if a critical point c is not floating then f can be slightly perturbed so that c becomes sinking; it is closely related to Lemmas 3.1 and 3.2 of [BM2].

Lemma 3.10. *Let $\varepsilon > 0$, a finite set $A \subset [0,1]$ and $f \in \mathcal{PN}$ be given. Suppose that a critical point c of f is not floating; moreover, if i is the greatest integer such that the point $f^i(c)$ is critical then $f^i(c)$ does not belong to A. Then there is a map $g \in \mathcal{PN}$ such that $\|f-g\|_2 < \varepsilon$, $f(x) = g(x)$ for every $x \in A$, the sets of critical points of f and g coincide and c is sinking for g.*

Proof. We may assume that c itself does not belong to A. Since c is not floating then it is either almost sinking or weakly sinking (if it is sinking, there is nothing to prove). If it is almost sinking then a small bump perturbation at c will push the orbit of c into the basin of attraction of a periodic sink. Here we use the assumption that c does not belong to A which allows us to make such perturbation. If this sink is attracting, we get c sinking. If it is neutral, we get c weakly sinking. Hence, from now on we can assume that c is weakly sinking. In this case a small perturbation in a neighborhood of a neutral periodic point x from the limit set of c makes x attracting (we can have $g(x) = f(x)$ in case $x \in A$) and leaves c attracted to the orbit of x, thus making c sinking. This completes the proof. ■

Before we prove our second preliminary result, we need a theorem similar to Theorem 1.3 of [BM2]. For the sake of convenience we write $[a;b]$ for the interval $[a,b]$ if $a<b$ and $[b,a]$ if $b<a$ (the same applies to other types of intervals).

Theorem 3.11. *For $f \in \mathcal{PS}$ the following properties hold.*

(1) *There are no wandering intervals.*
(2) *Points with finite limit sets are dense in* $[0, 1]$.
(3) *If there are no critical and no precritical points in* $(x; y)$ *then the limit set of every* $z \in [x; y]$ *is a periodic orbit.*
(4) *Any floating point is the limit from both sides (one side for* 0 *and* 1*) of both (pre)periodic and precritical points.*

Proof. (1) It is easy to check that $\mathcal{PS} \subset \mathcal{N}F^{1+bv}$, where the space $\mathcal{N}F^{1+bv}$ is defined in [MS, page 285]. The proof on non-existence of wandering intervals in Chapter IV of [MS] works for maps from $\mathcal{N}F^{1+bv}$, so (1) holds.

(2) Follows immediately from (1) and Lemma 6.1 of [B].

(3) Follows immediately from (1) and Lemma 3.1 of Chapter II of [MS].

(4) Assume that x is a floating point. Suppose that there is $y \neq x$ such that there are no precritical points in $(x; y)$. We may assume that there are also no critical points in $(x; y)$ (if there is a critical point in $(x; y)$ then we may replace y by the closest to x critical point in $(x; y)$). By (3) the limit set of x is a periodic orbit Q. Since x is floating, Q cannot be sinking or neutral, so it is repelling which easily implies that x itself is (pre)periodic. Now the fact that $(x; y)$ contains no precritical points implies that points from Q lie on the boundary of the basin of attraction of some periodic point of f. Therefore x is almost sinking, which contradicts the assumption that it is floating. This proves that x is the limit from both sides of precritical points.

Suppose now that there is $y \neq x$ such that there are no (pre)periodic points in $(x; y)$. By (2), there is $z \in (x; y)$ with a finite limit set. Since z is not (pre)periodic then it is easy to see that z is in the basin of attraction B of a periodic sink. Moreover, since the points (except perhaps $0, 1$) from the boundary of the basin of attraction of any cycle are periodic, we conclude that the whole interval $(x; z)$ is contained in B. Thus, x cannot be floating, a contradiction. This completes the proof of (4). ∎

Now we can state our second preliminary lemma. It basically sums up the results of Lemmas 3.3 and 3.4 of [BM2].

Lemma 3.12. *Let c be a floating non-recurrent critical point of $f \in \mathcal{PS}$. Then there is an arbitrarily small bump perturbation of f at c for which c is precritical.*

Proof. First assume that c has a finite limit set P. Then it is well known that P is a periodic orbit. Let $x \in P$; then x is a periodic repelling point not lying on the boundary of a basin of attraction of any periodic orbit of f (we rely upon the fact that c is floating). It is easy to see that then $f^j(c) = x$ for some j. Denote the period of x by n.

By Theorem 3.11 (4), there are precritical points arbitrarily close to x. There is a small neighborhood U of x such that a suitable branch of $h = (f|_U)^{-n}$ is a contraction and x is its fixed point. Choose a precritical point $y \in U$. Then $\lim_{k \to \infty} h^k(y) = x$. There is an arbitrarily small bump perturbation g of f at c such that $g^j(c) = h^k(y)$ for some k. If it is sufficiently small, the orbits of $g^j(c)$

under f and g coincide until they get to a critical point. Therefore c is precritical for g.

Now we consider the case when c is non-recurrent, but its limit set is infinite. By Theorem 3.11 (3), we can find a precritical point y so close to c that $y' = \tau_c(y)$ is well defined and the orbit of $f(c)$ is disjoint from $[y; y']$. There is a neighborhood U of c, contained in $[y; y']$, and such that the orbit of y misses U before it gets to a critical point.

Let V be a small neighborhood of $f(c)$. We want to prove that there is a precritical point $z \in V$ whose orbit misses U before it hits a critical point. If the union of images of V is disjoint from $[y; y']$ then this follows from Theorem 3.11 (3). If the union of images of V is not disjoint from $[y; y']$ then there is the smallest k such that $f^k(V)$ intersects $[y; y']$. We have $f^k(f(c)) \notin [y; y']$, so there is $z \in V$ such that either $f^k(z) = y$ or $f^k(z) = y'$. Then z is precritical and $f^i(z) \notin [y; y']$ for $i < k$. Thus, z is the point we were looking for.

Now, for every bump perturbation g of f at c such that the support of $g - f$ is contained in U, the point z found above is also precritical for g. Since V is arbitrarily small, we can choose g so that $g(c) = z$. This completes the proof. ∎

The next lemma follows easily from Lemmas 3.10, 3.12 and the definition of a locally best map.

Lemma 3.13. *Let $f \in \mathcal{PS}$ be a locally best map. Then all non-floating critical points of f are sinking and all floating critical points of f are recurrent.*

Proof. Let c be a non-floating critical point of f. Denote by A the set of all p-critical points of f. Then by Lemma 3.10 we can find a map g arbitrarily close to f which coincides with f on A and for which c is precritical (in other words, a small bump perturbation can be made in a way that does not spoil precriticality of any existing precritical critical point). This contradicts the definition of a locally best map. Similarly we can apply Lemma 3.12 thus showing that f cannot have floating non-recurrent critical points either. This completes the proof. ∎

We summarize the main results of this section in the following theorem.

Theorem 3.14. *For every $f \in \mathcal{N}_2$ there exists $\varepsilon > 0$ such that for any open ball X in $\mathcal{PN}$ with center f and radius smaller than ε there are maps $g \in X \cap \mathcal{PS}$ and $h \in X \cap \mathcal{N}_2$, best in X and such that*

(1) *$2\|g - f\|_2$ is smaller than the radius of X, all non-floating critical points of g are sinking, all floating critical points of g are recurrent, and a sufficiently small bump perturbation at any critical point of g does not change itineraries of the critical points of g,*
(2) *g and h have the same critical points, whenever a critical point is sinking for g, it is also sinking for h, and whenever a floating critical point is super persistently recurrent for g, it is also floating super persistently recurrent for h.*

Proof. If $\varepsilon > 0$ is sufficiently small then the existence of $g \in X \cap \mathcal{PS}$, best in X and satisfying (1) follows from Lemmas 3.4, 3.6 and 3.13. Existence of $h \in X \cap \mathcal{N}_2$, best in X and satisfying (2), follows from Lemma 3.5 and Theorem 2.9. In order to apply Theorem 2.9 we have to know additionally that g and h have no wandering intervals and that all floating super persistently recurrent critical points of g have infinite limit sets. However, the absence of wandering intervals follows from Theorem A, Chapter IV of [MS] (for h) and Theorem 3.11 (for g). A recurrent point with finite limit set has to be periodic, so if it is critical, it is sinking. This completes the proof. ■

4. Main Theorem

In this section we finish the proof of Main Theorem. In fact, from the point of view of dynamics, here we make the main step in the proof.

In the case of a map $f \in \mathcal{PN}$ (and thus of a map $f \in \mathcal{PS} \subset \mathcal{PN}$) there is a specific set $C \equiv C_f$ of all critical points together with all points of discontinuity of f''. Suppose that for some critical point c there exist arbitrarily small bump perturbations g of f at c such that the g-itinerary of c is distinct from its f-itinerary. Then we say that f is *unstable at c*. Now we are ready to state our main step.

Theorem 4.1. *Assume that c is a critical point of $f \in \mathcal{PS}$ which is recurrent but neither precritical nor C_f-super persistently recurrent. Then f is unstable at c.*

If c is precritical then clearly f is unstable at c. Therefore the assumption that c is not precritical is unnecessary in the above theorem. However, we include it, since we will be both proving and using Theorem 4.1 with this assumption.

It is rather easy to figure out how to complete the proof of Main Theorem using Theorem 4.1.

Proof of Main Theorem. Let f, g, h and X be as in Theorem 3.14. By the first part of that theorem, all critical points of g which are not sinking and not precritical, are recurrent and g is stable at them. By Theorem 4.1, g is super persistently recurrent at them. Hence, by Theorem 3.14 (2), every critical point of h is either sinking, or precritical, or floating super persistently recurrent. Note that if a precritical critical point is not sinking, then it is recurrent, so in view of Corollary 2.11 it is also super persistently recurrent. Since X is a ball centered at f with arbitrarily small radius and $h \in X \cap \mathcal{N}_2$, this proves Main Theorem. ■

To prove Theorem 4.1, we need distortion lemmas for piecewise negative Schwarzian maps. It is well known that these lemmas hold for negative Schwarzian maps in our sense. We start by stating Koebe Lemma for negative Schwarzian maps without critical points inside the domain. It has a part about the distortion and a part about lengths of intervals (that follows from the first one). In fact, we need only the second part. As in [BM2], we state it in the form most useful for us, and we state the first part in the form good for the proof of the second part. The first part in essentially the same form can be found for instance in [Br]. Moreover, it can be obtained easily from the convexity of the function $1/\sqrt{|h'|}$ by way of integration. We supply only a proof of the second part.

Koebe Lemma. *Let $h : [a,b] \to \mathbb{R}$ be a function with negative Schwarzian and such that $h' \neq 0$ on (a,b).*

(1) *Let $a < a' < b' < b$. Assume that $|h(a') - h(a)| \geq \delta |h(b') - h(a')|$ and $|h(b) - h(b')| \geq \delta |h(a') - h(b')|$. Then for every $x, y \in [a', b']$ we have $|h'(x)|/|h'(y)| \leq ((1+\delta)/\delta)^2$.*

(2) *Let $a < a'' < b'' < b$. Assume that $|h(b'') - h(a'')| \leq \omega |h(a'') - h(a)|$ and $|h(b'') - h(a'')| \leq \omega |h(b) - h(b'')|$. Then $b'' - a'' < 2\omega(3+2\omega)^2(a'' - a)$ and $b'' - a'' < 2\omega(3+2\omega)^2(b - b'')$. In particular, if $\omega \leq 1$, then $b'' - a'' < 50\omega(a'' - a)$ and $b'' - a'' < 50\omega(b - b'')$.*

Proof of (2). Without loss of generality, we may assume that h is increasing. Set $\widetilde{a} = h(a'') - (h(b'') - h(a''))/(2\omega)$ and $\widetilde{b} = h(b'') + (h(b'') - h(a''))/(2\omega)$, and then $a' = h^{-1}(\widetilde{a})$, $b' = h^{-1}(\widetilde{b})$. Then the assumptions of (1) are satisfied with $\delta = 1/(2+2\omega)$. By (1) and Mean Value Theorem we get

$$\frac{(h(a'') - h(a'))/(a'' - a')}{(h(b'') - h(a''))/(b'' - a'')} \leq \left(\frac{1+\delta}{\delta}\right)^2 = (3+2\omega)^2.$$

Since $h(b'') - h(a'') = 2\omega(h(a'') - h(a'))$ and $a'' - a' < a'' - a$, we get $b'' - a'' < 2\omega(3+2\omega)^2(a'' - a)$. Similarly, $b'' - a'' < 2\omega(3+2\omega)^2(b - b'')$. If $\omega \leq 1$ then $3 + 2\omega \leq 5$, and we get the required estimates. ■

In order to estimate distortion of iterates we also need to know what happens close to the critical points. Estimates of this type were used for instance by Lyubich in First Distortion Lemma of [L1]. By $|I|$ we denote the length of an interval I.

Lemma 4.2. *Let f be a piecewise monotone map of class C^1. Assume that in a neighborhood of every critical point c of f we have $A|x - c| \leq |f'(x)| \leq B|x - c|$ for some positive constants A, B. Then there exists a positive constant $D(f)$ such that if I, J are closed intervals with a common endpoint and disjoint interiors such that there are no critical points in the interior of $I \cup J$ and $|f(I)| \geq |f(J)|$ then $|I|/|J| \geq D(f)\sqrt{|f(I)|/|f(J)|}$.*

Proof. Suppose that such a constant does not exist. Then there are sequences of intervals I_n, J_n satisfying our assumptions, such that

$$\lim_{n\to\infty} \frac{|I_n|^2}{|f(I_n)|} \cdot \frac{|f(J_n)|}{|J_n|^2} = 0. \tag{4.1}$$

By passing to a subsequence, we may assume that the sequences of corresponding endpoints of I_n and J_n converge.

Suppose that the intervals I_n converge to a non-degenerate interval. Then the limit $\lim_{n\to\infty} |I_n|^2/|f(I_n)|$ exists and is non-zero. If the intervals J_n converge to a non-degenerate interval, then the limit $\lim_{n\to\infty} |f(J_n)|/|J_n|^2$ exists and is non-zero, so (4.1) does not hold. If the intervals J_n converge to a point that is not critical, then the limit $\lim_{n\to\infty} |f(J_n)|/|J_n|^2$ is infinite, so (4.1) also does not hold. If the intervals J_n converge to a critical point c, then for sufficiently large n from the assumption on f' we get by integration

$$\frac{A}{2}|J_n|\,\alpha(J_n) \le |f(J_n)| \le \frac{B}{2}|J_n|\,\alpha(J_n), \tag{4.2}$$

where $\alpha(J_n)$ is the sum of the distances of the endpoints of J_n from c. Since $\alpha(J_n) \ge |J_n|$, we get $|f(J_n)| \ge (A/2)|J_n|^2$, so (4.1) also does not hold.

Suppose now that the intervals I_n converge to a non-critical point x. Since $|f(J_n)| \le |f(I_n)|$, the intervals J_n also have to converge to x. Then the limits of $|f(I_n)|^2/|I_n|^2$ and $|f(J_n)|^2/|J_n|^2$ are both equal to $|f'(x)|^2$, so since $|f(I_n)| \ge |f(J_n)|$, (4.1) does not hold.

Suppose at last that the intervals I_n converge to a critical point c. Then so do the intervals J_n. Thus, for sufficiently large n (4.2) holds for J_n and for I_n replacing J_n. Therefore

$$\frac{|I_n|^2}{|f(I_n)|}\cdot\frac{|f(J_n)|}{|J_n|^2} \ge \frac{2}{B}\cdot\frac{|I_n|}{\alpha(I_n)}\cdot\frac{A}{2}\cdot\frac{\alpha(J_n)}{|J_n|} = \frac{A}{B}\cdot\frac{|I_n|}{|J_n|}\cdot\frac{\alpha(J_n)}{\alpha(I_n)}. \tag{4.3}$$

Moreover, since $|f(J_n)| \le |f(I_n)|$, we get $(A/2)|J_n|\,\alpha(J_n) \le (B/2)|I_n|\,\alpha(I_n)$, so

$$\frac{|I_n|}{|J_n|} \ge \frac{A}{B}\cdot\frac{\alpha(J_n)}{\alpha(I_n)}. \tag{4.4}$$

If v_n is the distance of the common endpoint of I_n and J_n from c then $\alpha(J_n) > v_n \ge (\alpha(I_n) - |I_n|)/2$, so

$$\frac{\alpha(J_n)}{\alpha(I_n)} > \frac{1}{2} - \frac{|I_n|}{2\alpha(I_n)}. \tag{4.5}$$

Since $\alpha(J_n) > |J_n|$, we get by (4.1) and (4.3) $\lim_{n\to\infty} |I_n|/\alpha(I_n) = 0$. Therefore by (4.5)

$$\liminf_{n\to\infty} \frac{\alpha(J_n)}{\alpha(I_n)} \ge \frac{1}{2}. \tag{4.6}$$

Thus, by (4.4),

$$\liminf_{n\to\infty} \frac{|I_n|}{|J_n|} \ge \frac{A}{2B}. \tag{4.7}$$

Now (4.3), (4.6) and (4.7) contradict (4.1). This completes the proof. ■

Let us remark that in general, if a critical point c of f has order l instead of 2, then a similar proof shows that if I and J are close to c then we get a similar inequality as in the above lemma, except that $\sqrt{|f(J)|/|f(J)|}$ is replaced by $(|f(J)|/|f(J)|)^{1/l}$.

The next lemma follows easily from Lemma 4.2. To state it, we need the following definition. If an interval I is contained in the interior of an interval T then T will be called a Δ-*crinoline* of I if the ratio of the length of each component of $T \setminus I$ to the length of I is at least Δ. The constant η was defined in Section 2.

Lemma 4.3. *For every $f \in \mathcal{PN}$ there is a positive constant $B(f)$ such that if I, I', T, T' are intervals such that T is a Δ-crinoline of I for some $\Delta \geq 1$, T' is a component of $f^{-1}(T)$ not containing 0 or 1, $|T'| < \eta$, and I' is a component of $f^{-1}(I)$ contained in T', then T' is a $B(f)\sqrt{\Delta}$-crinoline of I'.*

Proof. From the definition of η it follows that there exists a constant K such that if c is a critical point and an interval J of length less than η contains c then τ_c is defined on J and $|\tau_c'(x)| \leq K$ for all $x \in J$. Note that $K \geq 1$.

Let I, I', T, T' be intervals as above. If T' does not contain any critical point then it is mapped onto T in a monotone way and by Lemma 4.2 T' is a $D(f)\sqrt{\Delta}$-crinoline of I'.

Assume now that $T' = [a, b]$ contains one critical point c. We may assume that f has a local maximum at c. Then $T = [p, q]$ with $f(a) = f(b) = p$ and $p < f(c) \leq q$. Let $I' = [s, t]$. If I' contains c then $f(s) = f(t)$ is the left endpoint of I, and the right endpoint of I is to the right of $f(c)$ (or is equal to $f(c)$). Since T is a Δ-crinoline of I with $\Delta \geq 1$, we have $f(s) - p \geq |I| \geq f(c) - f(s)$. Therefore we may apply Lemma 4.2 and we get

$$\frac{s-a}{c-s} \geq D(f)\sqrt{\frac{f(s)-p}{f(c)-f(s)}} \geq D(f)\sqrt{\frac{f(s)-p}{|I|}} \geq D(f)\sqrt{\Delta}. \tag{4.8}$$

Since the modulus of the derivative of τ_c is bounded by K, we have $K(c-s) \geq t-c$, so $(K+1)(c-s) \geq |I'|$. Together with (4.8) this gives us $(s-a)/|I'| \geq (D(f)/(K+1))\sqrt{\Delta}$. Similarly we get $(b-t)/|I'| \geq (D(f)/(K+1))\sqrt{\Delta}$, so T' is a $(D(f)/(K+1))\sqrt{\Delta}$-crinoline of I'.

If now I' does not contain c then we may assume that it is contained in (a, c). Then we have $p < f(s) < f(t) < f(c)$ and $I = [f(s), f(t)]$. By the assumptions, $f(s) - p \geq \Delta|I|$, so by Lemma 4.2 $s - a \geq D(f)\sqrt{\Delta}|I'|$. Hence $b - t > b - \tau_c(s) \geq (s-a)/K \geq (D(f)/K)\sqrt{\Delta}|I'|$. Therefore T' is a $(D(f)/K)\sqrt{\Delta}$-crinoline of I'.

In all cases T' is a $(D(f)/(K+1))\sqrt{\Delta}$-crinoline of I', so the lemma holds with $B(f) = D(f)/(K+1)$. ■

Classical Koebe Lemma and Lemma 4.3 allow us to derive a version of Koebe Lemma for chains. To make this derivation simpler, we restate a part of Koebe Lemma.

Lemma 4.4. *Let $f \in \mathcal{PS}$, let $I \subset T$ be intervals such that there is no exceptional point inside T, and $f(T)$ is a Δ-crinoline of $f(I)$ for some $\Delta \geq 1$. Then T is a $\Delta/50$-crinoline of I.*

The following proposition follows easily from Lemmas 4.3 and 4.4.

Proposition 4.5. *For any $f \in \mathcal{PS}$ and a natural number ν there exists a function $\omega_{f,\nu}$ such that $\lim_{\Delta\to\infty} \omega_{f,\nu}(\Delta) = \infty$ and if chains $(G_i)_{i=0}^{l}$ of order ν or smaller and $(H_i)_{i=0}^{l}$ are such that $H_i \subset G_i$ for every i and G_l is a Δ-crinoline of H_l for some Δ such that $\omega_{f,\nu}(\Delta) \geq 1$, then G_0 is an $\omega_{f,\nu}(\Delta)$-crinoline of H_0.*

Proof. In the situation as above, we decompose our chain $(G_i)_{i=0}^{l}$ into $2\nu + 1$ (or less) pieces. Each piece corresponds either to f restricted to some G_i that contains one exceptional point or to an iterate f^j restricted to G_i such that there is no exceptional point of f^j in G_i (and then it has negative Schwarzian). We go back along the chain. In the first case we use Lemma 4.3, in the second case Koebe Lemma. During each step, if the last interval G_{i+j} of the piece was a Λ-crinoline of H_{i+j} then the first interval G_i of the piece is a $B(f)\sqrt{\Lambda}$-crinoline of H_i in the first case, and a $\Lambda/50$-crinoline of H_i in the second case.

Therefore the composition of the pull-backs corresponding to the second and the first case leads to the function $\varphi(t) = B(f)\sqrt{t/50}$ such that if to begin with G_s is a Λ-crinoline of H_s then after such two pull-backs G_r will be a $\varphi(\Lambda)$-crinoline of H_s. Since the order of the chain $(G_i)_{i=0}^{l}$ is at most ν we conclude that while pulling G_l and H_l back and estimating the distortion we will have to apply the function φ at most ν-times; other than that we may also need to apply a monotone negative Schwarzian pull back one more time. Denote by $\Phi = \varphi^\nu$ the ν-fold composition of the function φ. Then the function $\omega_{f,\nu} = \Phi/50$ satisfies the conditions of the proposition. ■

We apply Proposition 4.5 to prove the following sufficient condition for instability at a critical point. Let us remind the reader that the intervals $M_n(\cdot)$ and $H_n(\cdot)$ were defined between Lemma 2.3 and Theorem 2.4.

Proposition 4.6. *Let c be a non-precritical critical point of a map $f \in \mathcal{PS}$. Assume that there exists ν such that for every Δ there exist a positive integer k and a Δ-crinoline I of one of the parts into which $f^{k+1}(c)$ divides $M_k(f(c))$, such that the pull-back $\mathcal{G}$ of I along $f(c), \dots, f^{k+1}(c)$ has order ν or less. Then f is unstable at c.*

Proof. The point $f(c)$ divides $H_k(f(c))$ into two parts. Let J be the part whose image under f^k is the part of $M_k(f(c))$ from the statement of the proposition. Since f^k is monotone on J, there is a chain $(G_i)_{i=0}^{k}$ such that $f^i(J) \subset G_i$ for all i and $f^k(J) = G_k$. On the other hand, since $f^k(J) \subset I$, the intervals G_i are contained in the corresponding intervals of $\mathcal{G}$. Therefore we may use Proposition 4.5 if Δ is sufficiently large. We get that the first interval K of $\mathcal{G}$ is a Λ-crinoline of J, where $\Lambda = \omega_{f,\nu}(\Delta)$. By Proposition 4.5, we can get Λ as large as we want by choosing Δ sufficiently large.

We may assume that f has a local maximum at c. Let L be the component of $f^{-1}(K)$ containing c. Since K is smart, c is the only critical point in L. Therefore $L = [a, \tau_c(a)]$ for some a and $f(a)$ is the left endpoint of K.

Let us fix $\alpha \in (0,1)$. Let A be the set of all critical points of f^k in $[f(a), f(c)]$. For $x \in A$ we denote by $\xi(x)$ the smallest positive integer such that $f^{\xi(x)}(x)$ is a critical point. Clearly, $\xi(x) < k$ for every $x \in A$. Let us look at the rightmost point y of A. If $|J| < \alpha(f(c) - y)$ then we set $b = y$ and stop. Otherwise, we replace J by $[y, f(c)]$ and A by $\{x \in A : \xi(x) < \xi(y)\}$ and repeat the procedure. If at a certain moment the set replacing A is empty then we set $b = f(a)$. We end up with $b \in [f(a), f(c))$ and $m \le k$ such that f^m is monotone on $[b, f(c)]$ and either there is exactly one $b' \in (b, f(c))$ for which $f^m(b')$ is a critical point or $m = k$ and J is to the right of c. Since the order of $\mathcal{G}$ is smaller than or equal to ν, there are at most ν critical points of f^k in K. If Δ was so big that $\Lambda > \alpha^{-\nu}$ then in the first case $f(c)-b'<\alpha(f(c)-b)$. In the second case $|J|<\alpha(f(c)-b)$ and we define b' as the right endpoint of J.

Now, provided Δ is sufficiently large, we have the following situation. There are points $b\in[f(a),f(c))$ and $b'\in K$ and a positive integer m such that f^m is monotone on $[b,f(c)]$, the point $f^m(b')$ is critical for f and $|f(c)-b'|<\alpha(f(c)-b)$. Denote the preimage of b in (a,c) by p. Then the preimage of b in $(c,\tau_c(a))$ is $\tau_c(p)$. Set $\varepsilon=\min(c-p,\tau_c(p)-c)$ and $\delta=1000\cdot|b'-f(c)|/\varepsilon^2$. Let g be the (ε,δ)-bump perturbation of f at c, where the bump is added to f if $b'>f(c)$ and subtracted from f if $b'<f(c)$. Then $g(c)=b'$. Let us investigate the g-trajectory of b'.

We claim that the points $f^i(b')$, $i=1,2,\dots,m-1$, do not belong to $(p,\tau_c(p))$. Indeed, if $f^j(b')\in(p,\tau_c(p))$ then $f^{m-j-1}(z)$ is a critical point of f for $z=f^{j+1}(b')\in(b,f(c))$, contrary to the assumption that f^m is monotone on $[b,f(c)]$. This proves our claim. It follows by induction that $g^i(b')=f^i(b')$ for $i=1,2,\dots,m$. Hence, $g^{m+1}(c)=g^m(b')$ is a critical point of f. Since f and g have the same critical points, it is also a critical point of g. Since c was not precritical for f, the itineraries of c for f and g differ.

It remains to show that the perturbations we make can be made arbitrarily small. By Lemma 3.7, the C^2 norm of the perturbation described above is at most δ. Since the critical point c is non-degenerate, there is a constant $\gamma>0$ such that $f(c)-b\le\gamma\varepsilon^2$, independently of the choices we made. Then $\delta\le 1000\alpha\gamma$. We can make α as small as we want (by choosing Δ large enough), so the same applies to δ. This completes the proof. ∎

Proposition 4.6 allows us to continue the proof in the same way as in [BM2]. First we show that if we assume that Theorem 4.1 does not hold then we get a property similar to the reluctant recurrence of c, that allows us to use later our distortion estimates.

Lemma 4.7. *Assume that f and c satisfy the assumptions of Theorem 4.1, but f is stable at c. Then there is a number $\varepsilon>0$, an integer $\nu>0$ and a sequence $(n_i)_{i=1}^\infty$ of positive integers such that for every i the interval $[f^{n_i+1}(c)-\varepsilon,f^{n_i+1}(c)+\varepsilon]$ has a pull-back along $f(c),\dots,f^{n_i+1}(c)$ of order ν or less, such that f is monotone on every interval of this pull-back chain (except perhaps the last one).*

Proof. Since c is recurrent but not C_f-persistently recurrent, there is a number $\delta>0$, an integer $\nu>0$ and a sequence $(n_i)_{i=1}^\infty$ of positive integers such that for every i the interval $[f^{n_i+1}(c)-\delta,f^{n_i+1}(c)+\delta]$ has a pull-back along $f(c),\dots,f^{n_i+1}(c)$ of order ν or less. Since f is stable at c, by Proposition 4.6 there exists Δ such that for every i the interval $[f^{n_i+1}(c)-\delta,f^{n_i+1}(c)+\delta]$ is not a Δ-crinoline of any of the intervals into which $f^{n_i+1}(c)$ divides $M_{n_i}(f(c))$. This means that $M_{n_i}(f(c))$ contains $[f^{n_i+1}(c)-\varepsilon,f^{n_i+1}(c)+\varepsilon]$, where $\varepsilon=\delta/(2\Delta+1)$. Now the pull-backs of $[f^{n_i+1}(c)-\varepsilon,f^{n_i+1}(c)+\varepsilon]$ satisfy the conditions of the lemma. ∎

The following lemma is similar to Lemma 4.1 of [BM2].

Lemma 4.8. *Assume that f and c satisfy the assumptions of Theorem 4.1, but f is stable at c. Then c has infinite limit set and for any $\beta,\varepsilon_0>0$ there exist points*

a (close to c), b (close to $f(c)$), and a number n such that:

(1) *if f has local maximum at c then $f(a) < f(c) < b$, whereas if f has local minimum at c then $f(a) > f(c) > b$;*
(2) *$|a - c| < \varepsilon_0$ and $|\tau_c(a) - c| < \varepsilon_0$;*
(3) *$|b - f(c)| < \beta |f(c) - f(a)|$;*
(4) *$f^n|_{[f(a);b]}$ is monotone;*
(5) *the orbit of $f^n(b)$ misses the interval $[a; \tau_c(a)]$;*
(6) *if $f^i(b) \in [a; \tau_c(a)]$ for some i then $f^i(f(c))$ lies on the same side of c as $f^i(b)$, but farther away from c.*

Proof. Since c is critical but not precritical, it is not periodic. Therefore, since it is recurrent, it has infinite limit set.

For the sake of definiteness we may assume that f has local maximum at c. Let ε, ν and n_i be as in Lemma 4.7. Set $M_i = [f^{n_i+1}(c) - \varepsilon, f^{n_i+1}(c) + \varepsilon]$ and let H_i be the first interval of the pull-back chain from Lemma 4.7, whose last interval is M_i. Because of monotonicity, we have $f^{n_i}(H_i) = M_i$. We may assume that the maps $f^{n_i}|_{H_i}$ are either all increasing or all decreasing, $f^{n_i}(f(c)) \to y$ for some y and $[y - \varepsilon/2, y + \varepsilon/2] \subset M_i$ for every i (replace the sequence (n_i) by its subsequence if necessary).

Denote by H_i^- the part of H_i lying to the left of $f(c)$ and set $M_i^- = f^{n_i}(H_i^-)$. The reason why we are interested in H_i^- is that this set is the image of some neighborhood of c (since f has a local maximum at c).

All sets M_i^- lie on the same side of $f^{n_i}(f(c))$ and the points $f^{n_i}(f(c))$ approach y. Therefore every point z on one side of y does not belong to M_i^- for sufficiently large i (if the sets M_i^- lie to the left of $f^{n_i}(f(c))$ then we look at z to the right of y and vice versa). Let $\beta, \varepsilon_0 > 0$ be given (we may assume that β is small). Let us choose z as above, such that

$$|z - y| < \alpha, \tag{4.9}$$

where

$$\alpha = \frac{\varepsilon}{8\omega_{f,\nu}^{-1}(1/\beta)} \tag{4.10}$$

($\omega_{f,\nu}$ is the function from Proposition 4.5), and the limit set of z is finite. This is possible by Theorem 3.11 (2).

Now we give the precise meaning to the assumption that β is small. We need $\beta \le 1$, so that we can apply later Proposition 4.5. Moreover, we need β so small that $\alpha < \varepsilon/4$. This is possible, since $\lim_{t\to\infty} \omega_{f,\nu}(t) = \infty$.

Since $\alpha < \varepsilon/4 < \varepsilon/2$, we get $z \in [y - \varepsilon/2, y + \varepsilon/2] \subset M_i$ for every i.

Since the limit set of c is infinite, so is the limit set of $f(c)$. Thus, $f(c)$ is bounded away from the limit set of z. Moreover, by Theorem 3.11 (4), the length of H_i goes to 0 as $i \to \infty$. Thus, there exists j such that H_j is disjoint from the orbit of z, the component of the f-preimage of H_j containing c is shorter than ε_0, and

$$|f^{n_j}(f(c)) - y| < \alpha. \tag{4.11}$$

The component of the f-preimage of H_j containing c is of the form $[a;\tau_c(a)]$ for some a and then $H_j^- = [f(a), f(c)]$. Since this preimage is shorter than ε_0, (2) is satisfied. Let $n = n_j$ and let b be the $f^n|_{H_j}$-preimage of z. Then $[f(a);b] \subset H_j$, so (4) holds. Since M_j^- and z lie on the opposite sides of $f^n(f(c))$ and $f^n|_{H_j}$ is monotone, $f(a)$ and b lie on the opposite sides of $f(c)$. This proves (1). Since H_j is disjoint from the orbit of z, (5) holds.

Since $[y-\varepsilon/2, y+\varepsilon/2] \subset M_j$ and $\alpha < \varepsilon/4$, we get from (4.9) and (4.11) that each component of $M_j \setminus [f^n(f(c));z]$ has length at least $\varepsilon/4$. On the other hand, from (4.9) and (4.11) we get $|f^n(f(c)) - z| < 2\alpha$. Hence, by (4.10) M_j is an $\omega_{f,\nu}^{-1}(1/\beta)$-crinoline of $[f^n(f(c));z]$. Thus, from Proposition 4.5 we get (3).

Suppose that for some i we have $f^i(b) \in [a, \tau_c(a)]$. By (5), since $f^n(b) = z$, we get $i < n$. By (4) and (1), the points $f^i(f(a))$, $f^i(f(c))$ and $f^i(b)$ lie on the same side of c in this or reverse order. If $f^i(f(a))$ is the closest one to c among them, then either $[a;c]$ or $[\tau_c(a);c]$ is mapped into itself by f^{i+1} in a monotone way. Then the orbit of c is attracted to a periodic orbit, a contradiction. Thus, $f^i(b)$ is closer to c than $f^i(f(c))$, so (6) holds. ■

The proof of Theorem 4.1 is similar to the proofs of Lemmas 4.2 and 4.3 of [BM2].

Proof of Theorem 4.1. Assume that f and c satisfy the assumptions of Theorem 4.1, but f is stable at c. Then, by Lemma 4.8, c has infinite limit set and for any $\beta, \varepsilon_0 > 0$ there are a, b and n satisfying conditions (1)-(6) of Lemma 4.8. Since f is stable at c, there are $\delta_0, \varepsilon_0 > 0$ such that if $\varepsilon < \varepsilon_0$ and $\delta < \delta_0$ and g is an (ε,δ)-bump perturbation of f at c then the itineraries of c for f and g coincide.

Set $\beta = \delta_0/(500\gamma)$, where $\gamma = \sup_{x\in[0,1]} |f''(x)|$, $\varepsilon = \min(|a-c|, |\tau_c(a)-c|)$ and $U = (a;\tau_c(a))$. Then

$$(c-\varepsilon, c+\varepsilon) \subset U \tag{4.12}$$

and by Lemma 4.8 (2) we have $\varepsilon < \varepsilon_0$. Choose δ such that an (ε,δ)-bump perturbation g of f at c maps c to b. By (4.12), the support of $g - f$ is contained in U.

By Lemma 3.7 (1) we have $|b - f(c)| = s_{\varepsilon,\delta}(0) = \varepsilon^2\delta/1000$. Therefore, by Lemma 4.8 (3),

$$\delta = \frac{1000|b-f(c)|}{\varepsilon^2} < \frac{1000\beta|f(c)-f(a)|}{\varepsilon^2} = \frac{2\delta_0|f(c)-f(a)|}{\gamma\varepsilon^2}. \tag{4.13}$$

Since $f'(c) = 0$, we have $|f'(x)| \le \gamma|x-c|$ for any x, so $|f(c)-f(a)| \le \gamma\varepsilon^2/2$. Together with (4.13) this gives us $\delta < \delta_0$.

Set $z = f^n(b)$ and $m = n+1$. We claim that $f^m(c)$ and $g^m(c)$ lie on the opposite sides of z (non-strictly). Let us prove by induction that for any $j \le n$ the point $f^j(b)$ lies (non-strictly) between the points $f^j(f(c))$ and $g^j(b) = g^{j+1}(c)$. If $j = 0$ then $f^j(b) = g^j(b) = b$, so we have the induction base. Assume now that

$$f^i(b) \in [f^i(f(c)); g^i(b)] \tag{4.14}$$

for some $i < n$. Since the itineraries of c for f and g coincide, the interval $[f^i(f(c)); g^i(b)]$ belongs to one lap (the laps of f and g are the same). We have to prove that

$$f^{i+1}(b)) \in [f^{i+1}(f(c)); g^{i+1}(b)]. \tag{4.15}$$

We may assume that $f^i(f(c)) < g^i(b)$; the other case differs only by the direction of inequalities.

If $g^i(b) \notin U$ then $g^{i+1}(b) = f(g^i(b))$, and (4.15) follows from (4.14) and monotonicity of f on $[f^i(f(c)), g^i(b)]$. Assume now that $g^i(b) \in U$. By Lemma 4.8 (6), c lies on the same side of $f^i(f(c))$ as $f^i(b)$. By (4.14), we get

$$f^i(f(c)) < f^i(b) \le g^i(b) \le c. \tag{4.16}$$

Since f is monotone on $[f^i(f(c)), g^i(b)]$ and on U, it is monotone on the whole $[f^i(f(c)), c]$. If it is increasing, then f has a local maximum at c, so $g \ge f$ by Lemma 4.8 (1). Therefore from (4.16) we get $f^{i+1}(f(c)) < f^{i+1}(b) \le f(g^i(b)) \le g^{i+1}(b)$, and (4.15) follows. If it is decreasing, then f has a local minimum at c, so $g \le f$. Therefore from (4.16) we get $f^{i+1}(f(c)) > f^{i+1}(b) \ge f(g^i(b)) \ge g^{i+1}(b)$, and (4.15) also follows. This completes the induction step and proves the claim.

Set $x = f^m(c)$ and $y = g^m(c)$. Thus, x and y lie on the opposite sides of z. While y may coincide with z, we claim that x cannot. Indeed, by Lemma 4.8 (5) the f-orbit of $z = f^n(b)$ misses a neighborhood of c, so $c \notin \omega(z)$. On the other hand, c is recurrent and thus $c \in \omega(x)$. Hence $x \ne z$, so $x \ne y$.

Moreover, since the itineraries of c for f and g coincide, for each $k \ge 0$ the points $f^k(x)$ and $g^k(y)$ belong to the same lap (the laps of f and g coincide, too).

We claim that for every $i \ge 0$ we have $f^i(z) \in [f^i(x); g^i(y)]$ and $[f^i(x); g^i(y)]$ is non-degenerate. We prove it by induction. We know that for $i = 0$. Suppose that $f^j(z) \in [f^j(x); g^j(y)]$ and $[f^j(x); g^j(y)]$ is non-degenerate. Then the points $f^j(x), f^j(z), g^j(y)$ belong to the same lap on which either both f, g are increasing or they are both decreasing. Thus, application of f to the points $f^j(x)$ and $f^j(z)$ will change (or not) their order in the same way as application of g to the points $f^j(z) = g^j(z)$ and $g^j(y)$. On the other hand, $f(f^j(z)) = g(f^j(z))$ by the construction, which together with the previous remark shows that $f^{j+1}(z) \in [f^{i+j}(x); g^{i+j}(y)]$. Moreover, at least one of the intervals $[f^j(x); f^j(z)]$, $[f^j(z); g^j(y)]$ is non-degenerate, so at least one of the intervals $[f^{j+1}(x); f^{j+1}(z)]$, $[f^{j+1}(z); g^{j+1}(y)]$ is also non-degenerate. Therefore $[f^{i+j}(x); g^{i+j}(y)]$ is non-degenerate. This completes the induction step and proves the claim.

Thus, $f^i|_{[x,z]}$ is monotone for all i, which by Theorem 3.11 (4) implies that x is not floating. Hence, the f-limit set of $x = f^m(c)$ (and thus the f-limit set of c) is a periodic orbit, a contradiction. This completes the proof. ∎

References

[B] A. Blokh, *The spectral decomposition for one-dimensional maps*, Dynamics Reported **4** (1995), 1–59.

[BL1] A. Blokh and M. Lyubich, *Nonexistence of wandering intervals and structure of topological attractors of one-dimensional dynamical systems. II. The smooth case*, Ergod. Th. & Dynam. Sys. **9** (1989), 751–758.

[BL2] A. Blokh and M. Lyubich, *Measurable dynamics of S-unimodal maps of the interval*, Ann. Sci. Ecole Norm. Sup. (4), **24** (1991), 545–573.

[BM1] A. Blokh and M. Misiurewicz, *Collet-Eckmann maps are unstable*, Comm. Math. Phys. **191** (1998), 61–70.

[BM2] A. Blokh and M. Misiurewicz, *Dense set of negative Schwarzian maps whose critical points have minimal limit sets*, Discrete Contin. Dynam. Systems **4** (1997), 141–158.

[Br] H. Bruin, *Invariant measures of interval maps*, Ph. D. Thesis, Delft, 1994.

[J] M. V. Jakobson, *Smooth mappings of the circle into itself*, Mat. Sb. (N.S.) **85 (127)** (1971), 163-188. (Russian; English translation: Math. USSR - Sbornik, **14** (1971), 161-185)

[L1] M. Lyubich, *Nonexistence of wandering intervals and structure of topological attractors of one-dimensional dynamical systems. I. The case of negative Schwarzian derivative*, Erg. Th. & Dynam. Sys. **9** (1989), 737–749.

[L2] ———, *On the Lebesgue measure of the Julia set of a quadratic polynomial*, SUNY at Stony Brook, Preprint #1991/10 (1991).

[MS] W. de Melo and S. van Strien, *One-Dimensional Dynamics*, Springer Verlag, Berlin, 1993.

[MT] J. Milnor and W. Thurston, *On iterated maps of the interval*, Dynamical systems (College Park, MD, 1986–87), Lecture Notes in Math., vol. 1342, Springer, Berlin-New York, 1988, pp. 465–563.

Department of Mathematics, University of Alabama in Birmingham, University Station, Birmingham, AL 35294-2060

E-mail address: `ablokh@math.uab.edu`

Department of Mathematical Sciences, IUPUI, 402 N. Blackford Street, Indianapolis, IN 46202-3216

E-mail address: `mmisiure@math.iupui.edu`