DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume **32**, Number **2**, February **2012**

doi:10.3934/dcds.2012.32.433

pp. **433–466**

ON THE RELATIONS BETWEEN POSITIVE LYAPUNOV EXPONENTS, POSITIVE ENTROPY, AND SENSITIVITY FOR INTERVAL MAPS

Alejo Barrio Blaya and Víctor Jiménez López

Departamento de Matemáticas, Universidad de Murcia
Campus de Espinardo, 30100 Murcia, Spain

(Communicated by Lluis Alseda)

Abstract. Let $f : I = [0,1] \to I$ be a Borel measurable map and let μ be a probability measure on the Borel subsets of I. We consider three standard ways to cope with the idea of "observable chaos" for f with respect to the measure μ: $h_\mu(f) > 0$ —when μ is invariant—, $\mu(L^+(f)) > 0$ —when μ is absolutely continuous with respect to the Lebesgue measure—, and $\mu(S^\mu(f)) > 0$. Here $h_\mu(f)$, $L^+(f)$ and $S^\mu(f)$ denote, respectively, the metric entropy of f, the set of points with positive Lyapunov exponent, and the set of sensitive points to initial conditions with respect to μ.

It is well known that if $h_\mu(f) > 0$ or $\mu(L^+(f)) > 0$, then $\mu(S^\mu(f)) > 0$, and that (when μ is invariant and absolutely continuous) $h_\mu(f) > 0$ and $\mu(L^+(f)) > 0$ are equivalent properties. However, the available proofs in the literature require substantially stronger hypotheses than those strictly necessary. In this paper we revisit these notions and show that the above-mentioned results remain true in, essentially, the most general (reasonable) settings. In particular, we improve some previous results from [2], [6], and [23].

1. **Introduction.** Let (Ω, Σ) be a measurable space and let $f : \Omega \to \Omega$ be a measurable map. We are interested in studying the asymptotic behaviour of its *orbits* $(f^n(x))_{n=0}^\infty$, or, more precisely, in investigating when this behaviour becomes complicated (chaotic) for a large set of points x from the space Ω.

We use the word "large" in the measure-theoretic sense. Thus, in order to evaluate how large this set of points actually is, we must fix a probability measure μ on the σ-algebra Σ. It goes without saying that if we hope to arrive at some meaningful conclusions, then the measure μ should be meaningful as well.

Such is the case when μ is *invariant (for f)*, that is, $\mu(A) = \mu(f^{-1}A)$ for any $A \in \Sigma$, because the measure concentrates on points that are relevant from the dynamical point of view. A typical result: if Ω is a compact metric space X, and Σ is the σ-algebra $\mathcal{B} = \mathcal{B}(X)$ of Borel subsets of X, then the set of *recurrent* points of

2000 *Mathematics Subject Classification.* Primary: 37E05; Secondary: 37A05, 37A25, 37A35, 37D25, 37D45.

Key words and phrases. Absolutely continuous invariant measure, acip, invariant measure, Lyapunov exponents, metric entropy, Rohlin's formula, sensitivity to initial conditions.

This work has been partially supported by MEC (Ministerio de Educación y Ciencia, Spain), MICINN (Ministerio de Ciencia e Innovación, Spain) and FEDER (Fondo Europeo de Desarrollo Regional), grants MTM2005-03868 and MTM2008-03679/MTM, and Fundación Séneca (Comunidad Autónoma de la Región de Murcia, Spain), grants 00684/PI/04 and 08667/PI/08.

f (those points x belonging to the limit set of its orbit $(f^n(x))$) has full μ-measure (see, e.g., [27, Theorem 2.3, p. 29]).

Alternatively, μ could be required to be *absolutely continuous (with respect to the Lebesgue measure λ)*, that is, if $\lambda(A) = 0$, then $\mu(A) = 0$ for any $A \in \mathcal{B}$. Hence small sets in the usual sense cannot be too large in the sense of μ, which rather satisfies our geometric intuition. Of course, in order to properly speak about a "Lebesgue measure" we must restrict ourselves further, say to some Riemannian manifold M.

And finally, the nicest situation arises when μ is both invariant and absolutely continuous: then we call μ an *acip (for f)*.

Depending on the point of view we are interested in, one can apprehend this idea of "observable chaos" in a number of ways. According to the ergodic/probabilistic framework, the metric entropy $h_\mu(f)$ of f must be positive (here μ must be invariant with respect to f). The analytic approach requires that the set $L^+(f)$ of points with positive Lyapunov exponent has positive μ-measure; now we assume that μ is absolutely continuous with respect to λ. Finally, chaos is observable from a topological point of view whenever the set $S^\mu(f)$ of sensitive points to initial conditions with respect to μ has positive μ-measure. Let us investigate these notions in full detail.

In his 1958 paper [22] A. N. Kolmogorov, by introducing metric (or measure-theoretic) entropy, successfully adapted and developed Shannon's ideas on information theory to the setting of dynamical systems.

Let (Ω, Σ, μ) be a probability space and let $f : \Omega \to \Omega$ be measurable. The *metric entropy* of f, $h_\mu(f) \in [0, \infty]$, is defined by $h_\mu(f) = \sup h_\mu(f, \mathcal{A})$, with the supremum being taken over all finite measurable partitions $\mathcal{A}$. The precise definition of the number $h_\mu(f, \mathcal{A})$ is somewhat involved and we delay it until Subsection 3.2 but, informally speaking, it measures the average uncertainty, as n goes to ∞ and with respect to the partition $\mathcal{A}$, about the location of the n-iterate of a point under the action of the map f. Thus, in a sense, $h_\mu(f) > 0$ implies random, impossible to predict dynamics. Notice that in order to make sense of all of this one expects that the probability for the orbit of a point to visit one of the sets of the partition, after a given amount of time n, does not depend on n. In other words, μ must be invariant for f. (Incidentally, it is worth emphasizing that there is a way to characterize $h_\mu(f) > 0$ in terms of the existence of a set of points with "unpredictable" dynamics having positive μ-measure, very much in the fashion of our definitions of "analytic" and "topological" chaos below: see Remark 3.7.)

Lyapunov exponents are a classic analytic tool to measure chaos. They date back as early as around 1900, when J. Hadamard used them to prove the hyperbolicity of geodesic flows on manifolds of constant negative curvature.

Let M be a C^∞ compact manifold (possibly with boundary) and choose a Riemannian metric on $T(M)$. This means that we have fixed a scalar product (hence a norm $\|\cdot\|$) in every tangent space T_xM, $x \in M$, which depends on x in a differentiable way. The Riemannian metric induces in a canonical way a finite measure λ on the Borel subsets of M; after normalization, we can assume that λ is a probability measure. Since this measure apprehends the idea of "volume" for M, it is just natural to call it the *Lebesgue measure*; indeed, for the standard Riemannian metric in the cube I^n (throughout the paper, I stands for the unit interval $[0, 1]$) it is the usual Lebesgue measure. It must be emphasized that although different

Riemannian metrics induce different measures, all of them are equivalent: hence zero Lebesgue measure sets are uniquely defined.

Let $f : M \to M$ be a measurable map and let $x \in M$. If there are uniquely defined numbers

$$-\infty \leq \Lambda^{(1)}(x) < \Lambda^{(2)}(x) < \cdots < \Lambda^{(r(x))}(x) < \infty$$

and subspaces of T_xM

$$\{0\} = V^{(0)}(x) \subset V^{(1)}(x) \subset \cdots \subset V^{(r(x))}(x) = T_xM$$

such that

$$\lim_{n\to\infty} \frac{1}{n} \log \|df^n(x)u\| = \Lambda^{(i)}(x)$$

for all $u \in V^{(i)}(x) \setminus V^{(i-1)}(x)$, $i = 1, \ldots, r(x)$, then we call these numbers the *Lyapunov exponents* of f at x. Also we write $m_i(x) = \dim V^{(i)}(x) - \dim V^{(i-1)}(x)$ and define $\chi_f(x) = \sum_i (\Lambda^{(i)}(x))^+ m_i(x)$ (here $a^+ = \max\{a, 0\}$).

If M is the interval I, then everything is much simpler: there is just one Lyapunov exponent, $\Lambda_f(x)$, given by

$$\Lambda_f(x) = \lim_{n\to\infty} \frac{1}{n} \log |(f^n)'(x)|.$$

Notice that $\chi_f(x) = \Lambda_f^+(x)$ in this case.

Let $L^+(f) = \{x \in M : \chi_f(x) > 0\}$. If $x \in L^+(f)$, then its orbit and that of a point nearby diverge at a positive exponential rate. Hence it makes sense using the property $\lambda(L^+(f)) > 0$ (or, for that matter, $\mu(L^+(f)) > 0$ for some absolutely continuous measure μ with respect to the Lebesgue measure λ) as an indicator of "analytic chaos" for f.

Besides the probabilistic and analytic approaches previously described, sensitivity to initial conditions implements in a very appealing and natural way the idea of complicated behaviour from a topological point of view. Essentially, it is the formal translation to the discrete setting of what has become widely known as "the butterfly effect," famously introduced (in the framework of continuous-time models for climate prediction) by the meteorologist Lorenz in 1963 [24].

Definition 1.1. Let X be a compact metric space, let $f : X \to X$ be a Borel measurable map, and let μ be a probability measure on $\mathcal{B}$. Let $x \in X$ and $\delta > 0$. We say that x is *δ-sensitive (for f) with respect to μ*, written $x \in S_\delta^\mu(f)$ (or just *sensitive*, written $x \in S^\mu(f)$, if we do not need to put an emphasis on δ), if for every neighbourhood U of x there is some $n = n(x, U)$ such that the set $\{y \in U : \mathrm{dist}(f^n(x), f^n(y)) > \delta\}$ has positive μ-measure. If, moreover, $\mu(\{y \in U : \mathrm{dist}(f^m(x), f^m(y)) > \delta\}) > 0$ for every $m \geq n(x, U)$, then we say that x is *strongly δ-sensitive with respect to μ*.

We say that f has *sensitivity* (respectively, *full sensitivity*, *strong sensitivity*) *to initial conditions with respect to* μ if $\mu(S^\mu(f)) > 0$ (respectively, $S_\delta^\mu(f) = X$ for some $\delta > 0$, the set of δ-strongly sensitive points in the whole set X for some $\delta > 0$).

In the sequel we omit the words "with respect to μ" if μ is the Lebesgue measure.

Our definition of sensitivity is that originally given by Guckenheimer in 1979 in his seminal paper [16]. What we call "full sensitivity" is "sensitivity to initial conditions" as coined by Devaney in his 1986 book [13, Definition 8.2, p. 49]. The notion of strong sensitivity was first used (to our knowledge) by Blokh in 1982 [6] and brought to the fore recently [1].

To be precise our definitions are *stronger* than those just mentioned. The standard (*topological*) definition of sensitive point only requires that, for some $\delta > 0$, there are points y as close to x as required such that $\text{dist}(f^n(y), d^n(x)) > \delta$ for some (or all, in the case of strong sensitivity) sufficiently large n. The sets $S_\delta(f)$ and $S(f)$ of topological δ-sensitive and sensitive points are the natural ones to work with when no significant measure μ is available. However from this paper's perspective is quite natural to look for a definition of "observable sensitivity" that do not depend on zero measure modifications of the map f. In particular observe that if $\mu(A) = 0$ implies $\mu(f^{-n}(A)) = 0$ for every $n \geq 0$ (which is the case when μ is invariant, and also often the case when μ is absolutely continuous and f is reasonably smooth) and $g = f$ μ-a.e, then f is sensitive with respect to μ if and only if g is. We emphasize that since Definition 1.1 is more demanding than usual, our results below concerning sensitivity work as well for the topological versions of these notions. Also, notice if the *support* of μ (that is, the smallest closed subset of X having full measure, denoted $\operatorname{supp}\mu$) is the whole space X and every iterate of f is continuous μ-a.e, a point is δ-sensitive or δ-strongly sensitive with respect to μ if and only it is topologically δ-sensitive or topologically δ-strongly sensitive, respectively. By the way, it is worth emphasizing that even in such a setting (with $X = I$ and $\mu = \lambda$) there are examples of maps satisfying $S^\lambda(I) = S(f) = I$ but not having full sensitivity, and also maps having full but not strong sensitivity. On the other hand, if f is continuous and piecewise monotone, then $S(f) = I$ implies that f has strong sensitivity. See the appendix to this paper for the details.

(A brief notational interlude is required. Throughout the paper, when we say that a map $f : I \to I$ "*piecewise*" satisfies some property, then we mean that there are points $0 = a_0 < a_1 < \ldots < a_k = 1$ such that $f|_{(a_{i-1},a_i)}$ is continuous and satisfies this property for every $1 \leq i \leq k$. Notice that f need not be continuous at the points a_i.)

Presently we have considered metric entropy (in the setting of invariant measures), Lyapunov exponents (for absolutely continuous measures) and sensitivity (without any particular assumption on the measure μ). It must be emphasized that neither $h_\mu(f) > 0$ (when μ is invariant), nor $\mu(L^+(f)) > 0$ (when μ is absolutely continuous), nor $\mu(S^\mu(f)) > 0$ (even if μ is an acip) need necessarily imply "true complexity" for the map f: see Propositions 9.5-9.7. Of course none of these examples feature simultaneously the three types of chaos. However, it is often the case when one of them implies some other, and under the appropriate restrictions all of them amount to the same thing. The aim of this work is investigating and clarifying these connections as precisely as possible for interval maps. In doing this we improve a number of available results in the literature.

2. **Statement of the results.** It is well known that positive entropy implies sensitivity for continuous maps. More precisely, Blokh showed in [6] that if $f : I \to I$ is continuous, μ is invariant, and $h_\mu(f) > 0$, then $\mu(S(f)) > 0$. Later on Glasner and Weiss proved in [15] that if X is a compact metric space, $f : X \to X$ is continuous, μ is ergodic, $\operatorname{supp}\mu = X$, and $h_\mu(f) > 0$, then $S_\delta(f) = X$ for some $\delta > 0$. (We say that μ is *ergodic* if $f^{-1}(A) = A$ implies $\mu(A) = 0$ or $\mu(A) = 1$ for every measurable set A.) Actually, using Glasner and Weiss's result it is not hard to show that if f is continuous and $h_\mu(f) > 0$, then $\mu(S(f)) > 0$.

As our first result shows, the hypothesis of continuity can be completely disposed of. It must be emphasized that our argument is essentially the same as that employed by Katok in [19, Theorem 1.1]. Incidentally, Katok's idea was rediscovered recently by Cadre and Jacob in [11] to prove an alternate version of Theorem A(ii).

Theorem A. *Let X be a compact metric space, let $f : X \to X$ be a Borel measurable map and let μ be an invariant probability measure for f. Then the following statements hold:*

(i) If $h_\mu(f) > 0$, then f has sensitivity to initial conditions with respect to μ.
(ii) If $h_\mu(f) > 0$ and μ is ergodic, then $\mu(S^\mu_\delta(f)) = 1$ for some $\delta > 0$.
(iii) If $h_\mu(f) > 0$, μ is ergodic, and $\operatorname{supp} \mu = X$, then f has full sensitivity to initial conditions with respect to μ.

Somewhat surprisingly, the connections between positive Lyapunov exponents and sensitivity have not been explicitly investigated until quite recently, see [2]. Among other things, our next result shows that positive Lyapunov exponents imply sensitivity as far as absolutely continuous measures and piecewise monotone maps are concerned.

Theorem B. *Let $f : I \to I$ be piecewise monotone and let μ be an absolutely continuous probability measure for f. Assume $\mu(L^+(f)) > 0$. Then f has sensitivity to initial conditions with respect to μ.*

Additionally assume that $\mu(L^+(f)) = 1$, $\operatorname{supp} \mu = I$, and one of the following conditions is satisfied:

(a) f is piecewise C^2-diffeomorphic;
(b) μ is invariant for f;
(c) $f(x_+), f(x_-) \in E$ for any $x \in E$, where E is the set of discontinuity points and local extrema of f and $f(x_-)$ (respectively, $f(x_+)$) stands for the left-hand (respectively, right-hand) limit of f at the point x.

Then there is a number $\delta > 0$ such that if J is a subinterval of I then one of the components of $f^n(J)$ has length greater than δ for every n large enough. In particular, f has strong sensitivity to initial conditions with respect to μ.

Theorem B requires some comments. Concerning the first statement notice that continuity does not suffice. For instance the map f from Proposition 9.6 satisfies $\lambda(L^+(f)) = 1$ but $S(f) = \emptyset$.

Concerning the second statement, note that (c) is a kind of "Markov property" which, roughly speaking, implies that if f is not monotone on a certain interval (because it contains either a discontinuity point or a local extremum), then its iterates should not be too small (because they always contain a point from E, a finite set). We remark that strong sensitivity for f was proved in [2] assuming that (b), (c), and

(d) the density map of μ is bounded on some open subset of any given open set of I;
(e) $\log |f'|$ is μ-integrable;

were simultaneously satisfied. Hence our theorem substantially improves that of [2]. On the other hand, as the following result emphasizes, if neither of the conditions (a), (b) or (c) holds, then not even full sensitivity can be guaranteed.

Recall that we say that c is *critical* for a map f if f is differentiable at c and $f'(c) = 0$.

Theorem C. *There is a C^1-map g with just one critical point such that $\lambda(L^+(g)) = 1$ and it does not have full sensitivity to initial conditions.*

The relations between metric entropy and Lyapunov exponents have been investigated to a great depth. The essential of it: if μ is invariant, then $h_\mu(f) > 0$ implies $\mu(L^+(f)) > 0$; if μ is an acip, the converse statement is true as well.

Concerning the first result, the starting point is a theorem by Margulis (see [19]) who proved in 1968 that if $f : M \to M$ is a C^1-diffeomorfism and μ is invariant for f, then

$$h_\mu(f) \leq \int_M \chi_f \, d\mu.$$

Hence if f has positive metric entropy, then $\mu(L^+(f)) > 0$. In 1978 Ruelle extended the previous inequality (to which we refer in the sequel as the *Margulis-Ruelle inequality*) to arbitrary C^1-maps [32]. In fact the hypothesis on f can be somewhat relaxed without altering Ruelle's proof, see for instance Theorem 7.1. In [17] Hofbauer proved the inequality (for interval maps) under some hypotheses of piecewise monotonicity (here infinitely many pieces are allowed) and bounded variation derivative for f, also assuming the ergodicity of μ and $h_\mu(f) > 0$.

If μ is an acip, then it often happens

$$h_\mu(f) = \int_I \chi_f \, d\mu = \int_I \Lambda_f \, d\mu = \int_I \log |f'| \, d\mu,$$

which implies that $h_\mu(f) > 0$ and $\mu(L^+(f)) > 0$ are equivalent properties. The equality $h_\mu(f) = \int_I \log |f'| \, d\mu$ was proved by Rohlin in 1961 under relatively strong assumptions: f must be piecewise C^1-diffeomorphic, each of the C^1-pieces $f|_{(a_{i-1},a_i)}$ can be extended to a C^1-map in $[a_{i-1}, a_i]$, and derivatives have absolute value greater than 1 [31]. Ledrappier proved that the expanding condition is [23], so we call $h_\mu(f) = \int_I \log |f'| \, d\mu$ the *Rohlin-Ledrappier formula* in what follows.

Our next theorem shows that to prove the Rohlin-Ledrappier formula a piecewise Lipschitz condition suffices. Recall that a map $f : J \to \mathbb{R}$ is *Lipschitz* if there is a number $L > 0$ (a *Lipschitz constant for f*) satisfying $|f(y) - f(x)| \leq L|y - x|$ for every $x, y \in J$. We emphasize that when speaking about a piecewise Lipschitz map no monotonicity condition on its pieces of continuity is required.

Theorem D (the Rohlin-Ledrappier formula)**.** *Let $f : I \to I$ be a piecewise Lipschitz map and let μ be an acip for f. Then $0 \leq h_\mu(f) = \int_I \log |f'| \, d\mu = \int_I \Lambda_f \, d\mu < \infty$.*

In particular, $h_\mu(f) > 0$ if and only if $\mu(L^+(f)) > 0$.

The Rohlin-Ledrappier formula is a standard fact from interval ergodic theory but to outsiders it may not be so well known. For instance it could be used to prove the main result in [4] without a redundant hypothesis on the μ-integrability of $\log |f'|$. It must be stressed that Ledrappier's paper is very concise and difficult to follow to non-specialists on the ergodic side of dynamics, and according to our knowledge a complete, detailed and reader-friendly proof is not generally available. Besides generalizing it, our paper intends to render this important result accesible to a wider audience; in particular no serious knowledge of ergodic theory is expected. In fact, a very nice thing on the Rohlin-Ledrappier formula is that it effectively serves as a tutorial on this relevant topic: while checking the proof the reader will become acquainted with such key notions as, among others, Birkhoff's ergodic theorem, decomposition of an invariant measure into its ergodic components, and the Perron-Frobenius operator.

We see that under rather mild conditions on f and μ, positive metric entropy and $\mu(L^+(f)) > 0$ are equivalent and imply sensitivity to initial conditions with respect to μ. There exist counterexamples for the converse statements with very good differentiability properties (see Propositions 9.8-9.10). The last theorem of the paper shows that these examples cannot be stretched further because, in particular, it implies that if f is analytic and μ is an acip for f, then f cannot have sensitivity with respect to μ unless $h_\mu(f) > 0$.

In what follows, we say that a critical point c of f is *nonflat* if there is $n \geq 2$ such that f is of class C^{n+1} in a neighbourhood of c and $f^{(n)}(c) \neq 0$.

Theorem E. *Let $f : I \to I$ be a C^3-map with nonflat critical points and let μ be an acip for f. If $\mu(S(f)) > 0$, then $h_\mu(f) > 0$.*

Notice that, as a corollary, if $f : I \to I$ is a C^3-map with nonflat critical points and μ is an acip, then $h_\mu(f) > 0$, $\mu(L^+(f)) > 0$ and $\mu(S^\mu(f)) > 0$ are equivalent properties becoming, so to say, complementary pieces of evidence of the large, intrinsic dynamical complexity of the map f.

3. **Notation and some preliminary facts.** Before going to the proof of our theorems in the next sections we recall, and sometimes improve, some well-known results on a variety of topics.

3.1. **On partitions and conditional expectations.** A good general reference on the subject is [30, pp. 1–18].

Let (Ω, Σ) be a measurable space. When speaking about a *partition* of Ω we always refer to a finite partition of Ω into elements of Σ. If $\mathscr{A}$ and $\mathscr{B}$ are partitions of Ω, then we say that $\mathscr{A} \leq \mathscr{B}$ if for every $B \in \mathscr{B}$ there is some $A \in \mathscr{A}$ such that $B \subset A$. Similarly, if $\mathcal{C}$ and $\mathcal{D}$ are subσ-algebras of Σ, then $\mathcal{C} \leq \mathcal{D}$ means that $\mathcal{C} \subset \mathcal{D}$.

If $\{\mathscr{A}_n\}_{n=1}^m$ are partitions of Ω, then we denote by $\bigvee_{n=1}^m \mathscr{A}_n$ the partition consisting of the sets $A_1 \cap A_2 \cap \cdots \cap A_m$, with $A_n \in \mathscr{A}_n$ for every n. If $\{\mathcal{C}_n\}_{n=1}^m$ are subσ-algebras of Σ (here $m = \infty$ is possible), then $\bigvee_{n=1}^m \mathcal{C}_n$ denotes the smallest subσ-algebra of Σ containing all σ-algebras $\mathcal{C}_n$.

If $\mathscr{A}$ is a partition of Ω, then we use the same letter, but different type, to denote the (finite) subσ-algebra $\mathcal{A}$ of Σ generated by $\mathscr{A}$. Conversely, if $\mathcal{C}$ is a finite subσ-algebra of Σ, then there is exactly one partition $\mathscr{C}$ of Ω with the property that $\mathcal{C}$ is the subσ-algebra generated by $\mathscr{C}$. Notice in passing that $\mathscr{A} \leq \mathscr{B}$ if and only if $\mathcal{A} \leq \mathcal{B}$, and that the sub$\sigma$-algebra generated by $\bigvee_{n=1}^m \mathscr{A}_n$ is precisely $\bigvee_{n=1}^m \mathcal{A}_n$.

Let μ be a probability measure on Σ and let $\mathcal{C}$ be a subσ-algebra of Σ. The *conditional expectation operator given* $\mathcal{C}$, $E_\mu(\cdot|\mathcal{C})$, associates to every map $u \in L^1(\Omega, \Sigma, \mu)$ the unique μ-a.e. map $E_\mu(u|\mathcal{C}) \in L^1(\Omega, \mathcal{C}, \mu)$ satisfying

$$\int_C E_\mu(u|\mathcal{C})\, d\mu = \int_C u\, d\mu$$

for every $C \in \mathcal{C}$. In the particular case when $\mathcal{C}$ is a finite subσ-algebra of Σ and $\mathscr{C} = \{C_1, \ldots, C_p\}$ is the partition generating $\mathcal{C}$, we have

$$E_\mu(u|\mathcal{C}) = \sum_{j=1}^p 1_{C_j} \frac{\int_{C_j} u\, d\mu}{\mu(C_j)}.$$

It is easy to check that if $A \in \Sigma$, then $0 \leq E_\mu(1_A|\mathcal{C}) \leq 1$. Hence, if $\mathcal{A} = \{A_1, \ldots, A_k\}$ is a partition of Ω and $\mathcal{C}$ is a subσ-algebra of Σ, then the *conditional information of* $\mathcal{A}$ *given* $\mathcal{C}$, $I(\mathcal{A}|\mathcal{C})$, defined by

$$I(\mathcal{A}|\mathcal{C}) = \sum_{i=1}^{k} -1_{A_i} \log E_\mu(1_{A_i}|\mathcal{C}) = -\log \sum_{i=1}^{k} 1_{A_i} E_\mu(1_{A_i}|\mathcal{C}),$$

is nonnegative and we can define the *conditional entropy of* $\mathcal{A}$ *given* $\mathcal{C}$, $H_\mu(\mathcal{A}|\mathcal{C})$, by

$$H_\mu(\mathcal{A}|\mathcal{C}) = \int_\Omega I(\mathcal{A}|\mathcal{C})\, d\mu.$$

It turns out that $I(\mathcal{A}|\mathcal{C})$ is μ-integrable. Indeed we have $H_\mu(\mathcal{A}|\mathcal{C}) \in [0, \log k]$ because

$$\int_\Omega I(\mathcal{A}|\mathcal{C})\, d\mu = \int_\Omega \sum_{i=1}^{k} -E_\mu(1_{A_i}|\mathcal{C}) \log E_\mu(1_{A_i}|\mathcal{C})\, d\mu$$

(see [30, p. 10]), the convexity of the $x \log x$ map (which implies that if some numbers $x_i \geq 0$ satisfy $\sum_{i=1}^k x_i = 1$, then

$$\sum_{i=1}^{k} -x_i \log x_i \leq \log k, \tag{1}$$

see [35, Theorem 4.2, p. 79]), and the obvious fact that $\sum_{i=1}^k E_\mu(1_{A_i}|\mathcal{C}) = 1$.

If $\mathcal{C}$ is finite and $\mathscr{C}$ is its corresponding partition, then

$$H_\mu(\mathcal{A}|\mathcal{C}) = \sum_{i,j} -\mu(A_i \cap C_j) \log \frac{\mu(A_i \cap C_j)}{\mu(C_j)}.$$

In particular, the *entropy of the partition* $\mathcal{A}$, $H_\mu(\mathcal{A})$, is given by $H_\mu(\mathcal{A}) = H_\mu(\mathcal{A}|\mathcal{N})$, where $\mathcal{N}$ is the trivial σ-algebra $\{\emptyset, \Omega\}$, that is,

$$H_\mu(\mathcal{A}) = \sum_i -\mu(A_i) \log \mu(A_i).$$

Some useful results concerning the previous notions are given below:

Proposition 3.1. *Let* (Ω, Σ, μ) *be a probability space. Then the following statement holds:*

- *(i) If* $\mathcal{A}$, $\mathcal{B}$, *and* $\mathscr{C}$ *are partitions of* Ω, *then* $H_\mu(\mathcal{A} \vee \mathcal{B}|\mathcal{C}) = H_\mu(\mathcal{A}|\mathcal{C}) + H_\mu(\mathcal{B}|\mathcal{A} \vee \mathcal{C})$. *In particular,* $H_\mu(\mathcal{A} \vee \mathcal{B}) = H_\mu(\mathcal{A}) + H_\mu(\mathcal{B}|\mathcal{A})$.
- *(ii) If* $\mathcal{A}$ *is a partition of* Ω *and* $\mathcal{C}, \mathcal{D}$ *are sub*σ*-algebras of* Σ, *then* $\mathcal{C} \leq \mathcal{D}$ *implies* $H_\mu(\mathcal{A}|\mathcal{C}) \geq H_\mu(\mathcal{A}|\mathcal{D})$.
- *(iii) If* $\mathcal{A}$ *is a partition of* Ω *and* $(\mathcal{C}_n)$ *is an increasing sequence of sub*σ*-algebras of* Σ, *then* $(I(\mathcal{A}|\mathcal{C}_n))$ *converges to* $I(\mathcal{A}|\bigvee_n \mathcal{C}_n)$ *almost everywhere and* $(H(\mathcal{A}|\mathcal{C}_n))$ *converges to* $H(\mathcal{A}|\bigvee_n \mathcal{C}_n)$ *as* $n \to \infty$.

Proof. Properties (i) and (ii) are, respectively, [30, Corollary 1.11, p.8 and Remark 1.22, p.12]. Property (iii) is a consequence of the Doob martingale theorem [3, Theorem 7.6.2, p. 298], see also [30, Theorem 2.2(i), p. 16]. □

3.2. **Metric entropy.** Let (Ω, Σ, μ) be a probability space and assume that μ is invariant for a measurable map $f : \Omega \to \Omega$. Notice that if $\mathcal{A}$ is a partition of Ω or $\mathcal{C}$ is a subσ-algebra of Σ, and $n \geq 0$, then the obviously defined $f^{-n}(\mathcal{A})$ or $f^{-n}(\mathcal{C})$ is also a partition of Ω or a subσ-algebra of Σ, respectively. If $\mathcal{A}$ is a partition of Ω, then the nonnegative number

$$h_\mu(f, \mathcal{A}) = H_\mu \left(\mathcal{A} \,\middle|\, \bigvee_{n=1}^{\infty} f^{-n}(\mathcal{A}) \right) = \lim_{m \to \infty} \frac{1}{m} H_\mu \left(\bigvee_{n=0}^{m-1} f^{-n}(\mathcal{A}) \right)$$

is called the *entropy of f with respect to the partition $\mathcal{A}$* (for a proof of the second equality see, e.g, [30, pp. 24–25]).

The following important result, known as the *Shannon-McMillan-Breiman theorem*, provides an alternative way of computing the entropy of f with respect to a given partition. Later on it will prove instrumental to understand the connections between positive entropy and sensitivity. A good reference is [27, Theorem 1.2, p. 209].

Theorem 3.2. *Let (Ω, Σ, μ) be a probability space, let $f : \Omega \to \Omega$ be a measurable map, and assume that μ is invariant for f. Let $\mathcal{A}$ be a partition of Ω and, for every $x \in X$ and $n \geq 0$, denote by $A_n(x)$ the element of the partition $\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A})$ containing x. Then the limit*

$$h_\mu(f, \mathcal{A}, x) := \lim_{n \to \infty} -\frac{\log \mu(A_n(x))}{n}$$

exists for μ-a.e. $x \in \Omega$, the map $x \mapsto h_\mu(f, \mathcal{A}, x)$ is integrable, and

$$h_\mu(f, \mathcal{A}) = \int_X h_\mu(f, \mathcal{A}, x)\, d\mu.$$

Moreover if μ is ergodic, then $h_\mu(f, \mathcal{A}, x) = h_\mu(f, \mathcal{A})$ for μ-a.e. x.

Now the *(metric) entropy* of f is defined by $h_\mu(f) = \sup h_\mu(f, \mathcal{A})$, with the supremum being taken over all partitions $\mathcal{A}$.

A number of useful properties of entropy are listed below:

Proposition 3.3. *Let (Ω, Σ, μ) be a probability space, let $f : \Omega \to \Omega$ be a measurable map and assume that μ is invariant for f. Then the following statements hold:*

(i) $h_\mu(f^n) = n h_\mu(f)$ *for every* $n \geq 0$.

(ii) Let $B, C \in \Sigma$, $0 < \mu(B), \mu(C) < 1$. Assume that $f(B) \subset B$, $f(C) \subset C$, $\mu(B \cap C) = 0$ and $\mu(B \cup C) = 1$. Consider the invariant probability measures $\mu_B(A) = \mu(A \cap B)/\mu(B)$ and $\mu_C(A) = \mu(A \cap C)/\mu(C)$. Then $h_\mu(f) = \mu(B) h_{\mu_B}(f) + \mu(C) h_{\mu_C}(f)$.

(iii) If $h_\mu(f) = 0$, then $\mu(f^n(A)) = \mu(A)$ for every $n \in \mathbb{Z}$ and every $A \in \Sigma$.

If additionally Ω is a compact metric space X and $\Sigma = \mathcal{B}$, then the following statement holds:

(iv) Let $(\mathcal{P}_m)_{m=1}^{\infty}$ be an increasing sequence of partitions such that $\operatorname{diam} \mathcal{P}_m \to 0$ as $m \to \infty$. Then $h_\mu(f) = \lim_{m \to \infty} h_\mu(f, \mathcal{P}_m)$.

Proof. For (i) see [35, Theorem 4.13, p. 91]. For (ii) see [20, Corollary 4.3.17, p. 172]. Statement (iii) is [35, Corollary 4.14.2, p. 93]. Statement (iv) follows from the obvious fact $\mathcal{B} = \bigvee_{m=1}^{\infty} \mathcal{P}_m$ and [35, Theorem 4.22, p. 99]. □

3.3. **On ergodicity.** Birkhoff's ergodic theorem is the most important result of ergodic theory. For a proof see, e.g., [27, Theorem 1.1, pp. 89-90] or [35, Theorem 1.14, p. 34]. When applied to the map 1_A, and under the additional hypothesis of ergodicity for μ, it implies that for μ-a.e. point x the frequency with which the orbit of x hits the set A amounts to the measure of A, that is, time averages and space averages are exactly the same.

When we say below that a measurable map $u : X \to [-\infty, \infty]$ *has μ-integral*, then we mean that at least one of both integrals $\int_X u^+ \, d\mu$ or $\int_X u^- \, d\mu$ is finite (recall that $u^+ = \max\{u, 0\}$, while we denote $u^- = -\min\{u, 0\}$).

Theorem 3.4. *Let X be a compact metric space, let $f : X \to X$ be a Borel measurable map, and assume that μ is invariant for f. Let u have μ-integral. Then $\sum_{r=0}^{n-1} u(f^r(x))/n$ converges for μ-a.e. $x \in X$ to a map $u^*(x)$ having μ-integral and such that $\int_X u \, d\mu = \int_X u^* \, d\mu$. Moreover if μ is ergodic, then u^* is constant μ-a.e, that is,*

$$\int_X u \, d\mu = \lim_{n\to\infty} \frac{\sum_{r=0}^{n-1} u(f^r(x))}{n} \quad \text{for } \mu\text{-a.e. } x \in X.$$

In order to take most advantage of Birkhoff's theorem one needs ergodicity. A nice fact about this property is that it is often possible to assume it for an invariant measure without loss of generality. The reason behind this is that every invariant measure admits an *ergodic decomposition* [27, Theorem 6.4, p. 133].

Theorem 3.5. *Let X be a compact metric space, let $f : X \to X$ be a Borel measurable map, and assume that μ is invariant for f. Then, for μ-a.e. x, there are ergodic invariant probability measures μ_x having the following property: if u has μ-integral, then u has μ_x-integral for μ-a.e. $x \in X$, the map $x \mapsto \int_X u \, d\mu_x$ has μ-integral, and*

$$\int_X u \, d\mu = \int_X \left(\int_X u \, d\mu_x \right) \, d\mu.$$

Remark 3.6. It is worth emphasizing that in the proofs of Theorems 3.4 and 3.5 we mentioned earlier, it is assumed that u is μ-integrable. With this restriction Theorem 3.4 works for general probability spaces.

Concerning Theorem 3.5 our more general version easily follows from the standard one and the fact that any nonnegative measurable map is the increasing limit of nonnegative integrable maps and the monotone convergence theorem.

Theorem 3.4 requires a bit of care. Again we can assume that u is nonnegative. For every measure μ_x as in Theorem 3.5 we have that, regardless $\int_X u \, d\mu_x < \infty$ or $\int_X u \, d\mu_x = \infty$ (in the second case we use again a monotone convergence argument) $u^*(y) = \int_X u \, d\mu_x$ for μ_x-a.e. y; incidentally, one proves the special version of Theorem 3.4 when μ is ergodic in similar fashion. Then the set A of points at which u^* is well defined has full μ_x-measure for every x, hence (applying Theorem 3.5 to the map 1_A) has full measure μ and it makes sense to calculate $\int_X u^* \, d\mu$ even when $\int_X u \, d\mu = \infty$. Now two possibilities arise. If u is integrable, then $\int_X u \, d\mu = \int_X u^* \, d\mu$ is the standard version of Birkhoff's theorem. If $\int_X u \, d\mu = \infty$, then the standard monotone convergence trick implies $\int_X u^* \, d\mu = \infty$ as well.

Remark 3.7. From Theorem 3.2 and Proposition 3.3(iv) one can derive easily that, with the notation of Theorem 3.5, $h_\mu(f) = \int_X h_{\mu_x}(f) d\mu$. Hence $h_\mu(f) > 0$ if and only if the set of points x such that $h_{\mu_x}(f) > 0$ has positive μ-measure. Informally speaking, the measure μ_x concentrates on the points towards the orbit of x approaches more often, so $h_{\mu_x}(f) > 0$ means that the dynamics of f when restricted to these points is unpredictable enough. Notice that from this point of view a repelling fixed point x (while featuring positive Lyapunov exponent and topological sensitivity to initial conditions) does not contribute to the global complexity of f because μ_x concentrates on $\{x\}$, so trivially $h_{\mu_x}(f) = 0$.

3.4. **An equality related to the Perron-Frobenius operator.** A standard way to construct acips for interval maps is using the so-called Perron-Frobenius operator (see, e.g., [8, Chapters 4 and 5, pp. 75–109]).

Assume that $f : I \to I$ is *non-singular*, that is, $\lambda(A) = 0$ implies $\lambda(f^{-1}(A)) = 0$ for every $A \in \mathcal{B}$. The *Perron-Frobenius operator* P_f carries every map $u \in L^1(I, \mathcal{B}, \lambda)$ to the only map $P_f(u) \in L^1(I, \mathcal{B}, \lambda)$, λ-a.e., satisfying

$$\int_A P_f(u)\, d\lambda = \int_{f^{-1}(A)} u\, d\lambda$$

for every $A \in \mathcal{B}$. It turns out that an absolutely continuous probability measure μ is invariant for f if and only if its density ρ is a fixed point of the operator P_f, and sometimes it happens that P_f has nice contraction properties guaranteeing that if u is adequately chosen, then a subsequence of $(P_f^n(u))$ converges to this fixed point.

If f is piecewise diffeomorphic with bounded derivative, then there is an easy way to describe how the Perron-Frobenius works, namely

$$P_f(u)(z) = \sum_{\{y : f(y) = z\}} \frac{u(y)}{|f'(y)|} \quad \text{for } \lambda\text{-a.e. } z \in I$$

for every $u \in L^1(I, \mathcal{B}, \lambda)$ [8, pp. 85–86]. Thus, if ρ is the density of an acip μ for f, then we get

$$\rho(z) = \sum_{\{y : f(y) = z\}} \frac{\rho(y)}{|f'(y)|} \quad \text{for } \lambda\text{-a.e. } z \in I$$

and also, because f is non-singular,

$$\rho(f(x)) = \sum_{\{y : f(y) = f(x)\}} \frac{\rho(y)}{|f'(y)|} \quad \text{for } \lambda\text{-a.e. } x \in I. \tag{2}$$

In the rest of this subsection we extend (2) to the larger family of piecewise absolutely continuous maps. Recall that a map $f : J \to \mathbb{R}$ is *absolutely continuous* if for every $\epsilon > 0$ there is a number $\delta > 0$ such that $\sum_i (b_i - a_i) < \delta$ implies $\sum_i |f(b_i) - f(a_i)| < \epsilon$ whenever $\{(a_i, b_i)\}$ is a finite family of pairwise disjoint open subintervals of J. If f is absolutely continuous, then it is differentiable λ-a.e. (see for instance [12, Proposition 5.5.3(i), p. 153]) and carries zero measure sets to zero measure sets (this follows easily from the definition).

Also, recall that a set K is called *perfect* if it is closed and has no isolated points. Since every closed set is the union of a perfect set and a countable set we get:

Lemma 3.8. *If $C \subset I$ is closed, then there is a perfect set $K \subset C$ such that $\lambda(C) = \lambda(K)$.*

The following important result is due to Whitney [36]. Recall that $\operatorname{conv} C$ is the smallest interval containing C.

Theorem 3.9. *If f is differentiable λ-a.e, then for every $\epsilon > 0$ there are a closed subset $C \subset I$ and a C^1-map $g : \operatorname{conv} C \to \mathbb{R}$ such that $\lambda(I \setminus C) < \epsilon$ and $f(x) = g(x)$ for every $x \in C$.*

Lemma 3.10. *Let $f : I \to I$ be differentiable λ-a.e. and let*

$$B = \{x \in I : f' \text{ exists and } f'(x) \neq 0\}.$$

Then there are a family $\{K_j\}_{j=1}^{\infty}$ of pairwise disjoint perfect subsets of B and corresponding C_1-diffeomorphisms $\phi_j : \operatorname{conv} K_j \to \operatorname{conv} f(K_j)$ having the following properties:

(i) $\lambda(\bigcup_j K_j) = \lambda(B)$;
(ii) *every ϕ_j extends $f|_{K_j}$, that is, $\phi_j(x) = f(x)$ for every $x \in K_j$.*

Proof. If $\lambda(B) = 0$, then there is nothing to prove. If $\lambda(B) > 0$, then there are a closed set $C \subset B$ with $\lambda(C) > \lambda(B)/2$ and a C^1-map $g : \operatorname{conv} C \to \mathbb{R}$ satisfying $g(x) = f(x)$ for every $x \in C$ (Theorem 3.9). In view of Lemma 3.8, there is no loss of generality in assuming that C is perfect. Since $C \subset B$, this also implies $g'(x) = f'(x) \neq 0$ for every $x \in C$. Since g is C^1, there is a family of pairwise disjoint closed intervals $\{I_j\}_{j=1}^{r}$ such that $g'|_{\bigcup_j I_j}$ never vanishes and the set of points $x \in I \setminus \bigcup_j I_j$ with $g'(x) \neq 0$ has measure less than $\lambda(C) - \lambda(B)/2$. Additionally we can assume that none of the endpoints of the intervals I_j is a one-sided isolated point of C. Then the sets $K_j = C \cap I_j$, $1 \leq j \leq r$, are pairwise disjoint perfect subsets of B satisfying $\lambda(\bigcup_j K_j) > \lambda(B)/2$, and g extends each restriction $f|_{K_j}$ to a C^1-diffeomorphism between $\operatorname{conv} K_j$ and $\operatorname{conv} f(K_j)$.

We conclude the proof of the lemma by starting from $B \setminus \bigcup_{j=1}^{r} K_j$ and iterating the process. □

Lemma 3.11. *Let $f : I \to I$ be piecewise absolutely continuous and let $\{K_j\}$ be the sets from Lemma 3.10. Then the following statements hold:*

(i) *if $\lambda(N) = 0$, then $\lambda(f^{-1}(N) \cap \bigcup_j K_j) = 0$;*
(ii) $\lambda(f(I \setminus \bigcup_j K_j)) = 0$;
(iii) *if μ is an acip for f, then $\mu(\bigcup_j K_j) = 1$;*
(iv) $f^{-1}(\{f(x)\}) \subset \bigcup_j K_j$ *for λ-a.e. $x \in \bigcup_j K_j$.*

Proof. To prove (i) it suffices to show $\lambda(f^{-1}(N) \cap K_j) = 0$ for every j, which follows from $f^{-1}(N) \cap K_j = \phi_j^{-1}(N \cap \phi_j(K_j))$ for the diffeomorphism ϕ_j from Lemma 3.10(ii).

Now we prove (ii). Let B be defined as in Lemma 3.10 and write $N_0 = B \setminus \bigcup_j K_j$, $N_1 = \{x \in I : f'(x) \text{ does not exist}\}$, and $N_2 = \{x \in I : f'(x) = 0\}$. We have $\lambda(N_0) = 0$ by Lemma 3.10(i) and $\lambda(N_1) = 0$ because f is differentiable λ-a.e. Moreover, $\lambda(f(N_0)) = \lambda(f(N_1)) = 0$ because f is piecewise absolutely continuous. Finally, we use Lemma 3.8 and Theorem 3.9 to find perfect sets $C_k \subset N_2$ and C^1-maps g_k with $\lambda(\bigcup_k C_k) = \lambda(N_2)$ and $g_k(C_k) = f(C_k)$. Since $g_k'(x) = f'(x) = 0$ for every $x \in N_2$ and g_k is a C^1-map, we have $\lambda(f(C_k)) = \lambda(g_k(C_k)) = 0$ for every k. Therefore $\lambda(f(N_2)) = 0$. Since $f(I \setminus \bigcup_j K_j) = f(N_0) \cup f(N_1) \cup f(N_2)$, (ii) follows.

Assume that μ is an acip. We have just shown that $\lambda(f(I \setminus \bigcup_j K_j)) = 0$. Then $\mu(f(I \setminus \bigcup_j K_j)) = 0$ because μ is absolutely continuous and $\mu(I \setminus \bigcup_j K_j) \leq \mu(f^{-1}(f(I \setminus \bigcup_j K_j))) = 0$ because μ is invariant. We have proved (iii).

Finally we prove (iv). If a point x is given, then $f^{-1}(\{f(x)\}) \subset \bigcup_j K_j$ if and only if $x \notin f^{-1}(f(I \setminus \bigcup_j K_j))$. Thus we must prove $\lambda(f^{-1}(f(I \setminus \bigcup_j K_j)) \cap \bigcup_j K_j) = 0$, which follows from (i) and (ii). □

Lemma 3.12. *Let $f : I \to I$ be piecewise absolutely continuous and μ and acip for f. Let ρ be the density of μ. Then $\rho(f(x)) = \sum_{\{y \in \bigcup_j K_j : f(y)=f(x)\}} \rho(y)/|f'(y)|$ for λ-a.e. $x \in \bigcup_j K_j$.*

Proof. Let $P \subset I$ be a measurable set. Then

$$\begin{aligned}
\int_P \rho \, d\lambda &= \mu(P) = \mu(f^{-1}(P)) = \int_{f^{-1}(P)} \rho \, d\lambda \\
&= \sum_j \int_{K_j \cap f^{-1}(P)} \rho \, d\lambda + \int_{(I \setminus \bigcup_j K_j) \cap f^{-1}(P)} \rho \, d\lambda \\
&= \sum_j \int_{K_j \cap f^{-1}(P)} \rho \, d\lambda = \sum_j \int_{K_j \cap \phi_j^{-1}(P)} \rho \, d\lambda \\
&= \sum_j \int_{\phi_j(K_j) \cap P} \frac{\rho}{|\phi_j'|} \circ (\phi_j|_{K_j})^{-1} \, d\lambda \\
&= \int_P \sum_j 1_{\phi_j(K_j)} \cdot \left(\frac{\rho}{|\phi_j'|} \circ (\phi_j|_{K_j})^{-1} \right) d\lambda \\
&= \int_P \sum_j 1_{f(K_j)} \cdot \left(\frac{\rho}{|f'|} \circ (f|_{K_j})^{-1} \right) d\lambda;
\end{aligned}$$

we have used that μ is an acip, Lemma 3.11(iii), and the change of variables theorem. Recall also that $f'(x) = \phi_j'(x)$ for λ-a.e. $x \in K_j$ because K_j is perfect.

Since the set P was arbitrarily chosen, we see that $\rho = \sum_j 1_{f(K_j)} \cdot ((\rho/|f'|) \circ (f|_{K_j})^{-1})$ λ-a.e, that is,

$$\rho(z) = \sum_{\{y \in \bigcup_j K_j : f(y)=z\}} \frac{\rho(y)}{|f'(y)|} \quad \text{for } \lambda\text{-a.e. } z \in I.$$

Then

$$\rho(f(x)) = \sum_{\{y \in \bigcup_j K_j : f(y)=f(x)\}} \frac{\rho(y)}{|f'(y)|} \quad \text{for } \lambda\text{-a.e } x \in \bigcup_j K_j$$

by Lemma 3.11(i). □

Putting together Lemmas 3.10, 3.11(ii), 3.11(ii)(iv), and 3.12 we conclude:

Theorem 3.13. *Let f be a piecewise absolutely continuous map, let μ be an acip for f and let ρ be the density of μ. Then there are a family $\{K_j\}_{j=1}^{\infty}$ of pairwise disjoint perfect sets and corresponding C_1-diffeomorphisms $\phi_j : \operatorname{conv} K_j \to \operatorname{conv} f(K_j)$ having the following properties:*

- *(i)* $\lambda(f(I \setminus \bigcup_j K_j)) = 0$;
- *(ii)* *every ϕ_j extends $f|_{K_j}$;*
- *(iii)* $f^{-1}(\{f(x)\}) \subset \bigcup_j K_j$ *and* $\rho(f(x)) = \sum_{\{y : f(y)=f(x)\}} \rho(y)/|f'(y)|$ *for λ-a.e. $x \in \bigcup_j K_j$.*

3.5. **Negative Schwarzian derivative.** If J, K are intervals, $f : J \to K$ is a C^3-map, and $x \in J$ is not a critical point of f, then the *Schwarzian derivative of f at x*, $Sf(x)$, is defined by

$$Sf(x) = \frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2.$$

Notice that we may use the symbol "S" to denote both the Schwarzian derivative of a map and its set of sensitive points. Its correct meaning will always be clearly inferred from the context.

Maps with negative Schwarzian derivative are very important in one-dimensional dynamics. The reason is the following. If two diffeomorphisms $f : J \to K$ and $g : K \to L$ have negative Schwarzian derivative, then $S(g \circ f) < 0$ (see, e.g., [13, Proposition 11.3, p. 69]). On the other hand if $Sf < 0$, then the well-known Koebe lemma allows us to estimate its distortion regardless which map f exactly is. Combining both facts, if $f : I \to I$ and $Sf < 0$ outside the critical points of f, then one can often get very good control of the "nonlinearity" of *simultaneously* all iterates of f.

Next we formulate a version of Koebe's lemma (see [7, Lemma 3.4]) that suits our purposes. If $\delta > 0$ and K is a subinterval of an interval J with the property that both components of $J \setminus K$ have at least length $\delta\lambda(J)$, then we call J a *δ-scaled neighbourhood* of K.

Lemma 3.14. *If $\varphi : J \to \mathbb{R}$ has negative Schwarzian derivative and $K \subset J$ is an interval such that $\varphi(J)$ is a δ-scaled neighbourhood of $\varphi(K)$, then*

$$\lambda(A)/\lambda(K) \geq \left(\frac{\delta}{1+\delta}\right)^2 \lambda(\varphi(A))/\lambda(\varphi(K))$$

for every measurable set $A \subset K$.

The following lemma allows us to construct maps with negative Schwarzian derivative and prescribed (usual) derivatives at the endpoints of the interval. In what follows $\langle a, b\rangle$ denotes the smallest connected set containing $\{a, b\}$.

Lemma 3.15. *Let $J = \langle a, b\rangle$ and $K = \langle c, d\rangle$ be compact intervals. Let $u, v \in \mathbb{R}$ and assume that the numbers u, v, and $e := (d-c)/(b-a)$ have the same sign and $|u|, |v| < |e|$. Then there is a diffeomorphism $\varphi : J \to K$ having the following properties:*

(i) $\varphi(a) = c$, $\varphi(b) = d$;
(ii) $\varphi'(a) = u$, $\varphi'(b) = v$;
(iii) $\min\{|u|, |v|\} \leq |\varphi'(x)| \leq |e| + (2|e| - |u| - |v|)/3$ *for every* $x \in J$;
(iv) $S\varphi < 0$.

Proof. Clearly, there is no loss of generality in assuming $a = c = 0$, $b = d = 1$, and $0 < u, v < 1$. Next we show that the restriction of $\varphi(x) = (v + u - 2)x^3 + (3 - v - 2u)x^2 + ux$ to I does the job.

A routine calculation shows that φ is a diffeomorphism on I and that (i) and (ii) are satisfied. Moreover, the equation $\varphi'(x) = 0$ has two distinct real roots and hence, according to [13, Proposition 11.2, p. 69], φ has negative Schwarzian derivative outside its critical points, in particular in I. Finally, it can be checked that

$$\min_{x \in I} \varphi'(x) = \min\{u, v\},$$

and

$$\begin{aligned}\max_{x\in I}\varphi'(x) &= u+\frac{2-u-v}{3}+\frac{2(1-u)}{3}+\frac{(1-u)^2}{3(2-u-v)}\\ &< u+\frac{2-u-v}{3}+\frac{2(1-u)}{3}+\frac{(1-u)^2}{3(1-u)}\\ &= 1+\frac{2-u-v}{3},\end{aligned}$$

from which (iii) follows. □

4. **Positive entropy implies sensitivity.** In this section we demonstrate Theorem A.

First we prove (i). To this aim we use [35, Lemma 8.5, p. 187] to find a partition $\mathcal{A}=\{P_1,P_2,\dots,P_k\}$ of X such that $\operatorname{diam}\mathcal{A}$ is very small and having the additional property that the union set N of the boundaries of the sets P_i has zero measure. Due to Proposition 3.3(iv) we can assume additionally $h_\mu(f,\mathcal{A})>0$.

According to Theorem 3.2 there are a number $\theta>1$ and a set C such that $\mu(C)>0$ and $h_\mu(f,\mathcal{A},x)>\log\theta$ for every $x\in C$. Now, with the notation of Theorem 3.2, we can find a set $C'\subset C$ whose measure is very close to that of C (in particular $\mu(C')>0$) and a number n_0 such that $-\log\mu(A_n(x))/n>\log\theta$, and hence

$$\mu(A_n(x))<\frac{1}{\theta^n}, \tag{3}$$

for every $x\in C'$ and $n\geq n_0$.

Let $\epsilon>0$ be small enough so that $k^\epsilon<\theta$. If $\delta>0$ is also very small, then the set of points $M=\{x\in X:\operatorname{dist}(x,N)\leq\delta\}$ has measure less than $\epsilon\mu(C')$. Now Theorem 3.4, when applied to the map $u=1_M$, implies that there are a subset C'' of C' of positive measure and a number n_1 such that

$$\operatorname{Card}\{0\leq i<n: f^i(x)\in M\}<\epsilon n \tag{4}$$

for every $x\in C''$ and $n\geq n_1$.

We claim that $\mu(C''\cap S^\mu_\delta(f))=\mu(C'')>0$, which finishes the proof. By way of contradiction we assume that $D\cap S^\mu_\delta(f)=\emptyset$ for a subset D of C'' of positive measure. Then, for every $x\in D$, there exist a neighbourhood $U(x)$ of x and $U'(x)\subset U(x)$ with $\mu(U'(x))=\mu(U(x))$ such that $\operatorname{dist}(f^n(x),f^n(y))\leq\delta$ for every $n\geq 0$ and $y\in U'(x)$. Since Borel subsets of X can be approached in measure from below by compact sets we can suppose that D is compact, which allows us to find a finite cover $U(x_1),\dots,U(x_R)$ of D (hence $U'(x_1),\dots,U'(x_R)$ cover a subset D' of D such that $\mu(D')=\mu(D)$).

We define

$$A_\alpha=\{x\in I: f^i(x)\in P_{\alpha_i}, 0\leq i<n\}$$

for every finite sequence $\alpha=(\alpha_0,\dots,\alpha_{n-1})\in\{1,\dots,k\}^n$. Let $n\geq n_1$. If $y\in U'(x_r)$ for some $1\leq r\leq R$ and $0\leq i<n$, then either $f^i(y)$ and $f^i(x_r)$ belong to the same set of the partition $\mathcal{A}$ or $f^i(x_r)\in M$. Since the number of indexes i for which the second alternative holds is bounded by ϵn (4), we see that for every number r the family of sequences $\alpha\in\{1,2,\dots,k\}^n$ with the property that A_α intersects $U'(x_r)$ has cardinality at most $k^{\epsilon n}$. Hence the family $\mathcal{F}_n$ of sequences $\alpha\in\{1,\dots,k\}^n$ such that A_α intersects D' has cardinality at most $Rk^{\epsilon n}$.

Notice that if $x \in A_\alpha$ for some $\alpha \in \{1, \dots, k\}^n$, then $A_\alpha = A_n(x)$. Since $D' \subset D \subset C'' \subset C'$, (3) implies

$$\mu(D') \le \sum_{\alpha \in \mathcal{F}_n} \mu(A_\alpha) \le \frac{Rk^{\epsilon n}}{\theta^n}$$

for every $n \ge \max\{n_0, n_1\}$. Thus $\mu(D) = \mu(D') = 0$ and we have arrived at the desired contradiction.

The proof of (ii) involves no significant changes. In this case, if the number $0 < \kappa < 1$ is given, then the ergodic versions of Theorems 3.4 and Theorem 3.2 allow to get $\mu(C'') > 1 - \kappa$ and hence $\mu(C'' \cap S_\delta^\mu(f)) = \mu(C'') > 1 - \kappa$ as before. Since $\kappa > 0$ was arbitrarily chosen, we get $\mu(S_\delta^\mu(f)) = 1$.

Finally we prove (iii). We know that $\mu(S_\delta^\mu(f)) = 1$ for some $\delta > 0$ by (ii). Assume that $x \in X$ does not belong to $S_{\delta/2}^\mu(f)$. Then there are an open neighbourhood U of x and a full measure subset U' of U such that $\operatorname{dist}(f^n(x), f^n(z)) \le \delta/2$ for every $n \ge 0$ and $z \in U'$. Since $\operatorname{supp} \mu = I$ and $\mu(S_\delta^\mu(f)) = 1$, $U' \cap S_\delta^\mu(f)$ has positive measure (hence it is nonempty). Take $y \in U' \cap S_\delta^\mu(f)$ and a find a neighbourhood V of y contained in U. Let $z \in V' = V \cap U'$. Then $\operatorname{dist}(f^n(y), f^n(z)) \le \operatorname{dist}(f^n(y), f^n(x)) + \operatorname{dist}(f^n(x), f^n(z)) \le \delta/2 + \delta/2 = \delta$ for every n and $z \in V'$, which contradicts $y \in S_\delta^\mu(f)$.

5. **Positive Lyapunov exponents imply sensitivity.** We devote this section to prove Theorem B. Before doing so, two preparatory lemmas are required.

Let $f : I \to I$ be piecewise monotone. If $P : 0 = a_0 < a_1 < \cdots < a_k = 1$ is a finite set of points such that $f|_{(a_{i-1}, a_i)}$ is monotone for every i, then we call P a *singular set (for f)*. We emphasize that P need not be maximal with the required property.

We say that an open interval $J \subset I$ is *substantial* for a measure μ if $J \cap \operatorname{supp} \mu$ accumulates at both endpoints of J.

Lemma 5.1. *Let $f : I \to I$ be piecewise monotone, let P be a singular set, and let μ be an absolutely continuous probability measure. Then, for μ-a.e. $x \in L^+(f)$, there is a number $\delta = \delta(x) > 0$ with the following property: if U is a neighbourhood of x, then there are a substantial interval $J \subset U$ and an integer n such that $f^n|_J$ is continuous and monotone and $\lambda(f^n(J)) > \delta$.*

Proof. For every pair m, l of positive integers let $L_{m,l}$ be the set of points x satisfying $|(f^n)'(x)| > (1 + 1/m)^n$ for every $n \ge l$. Then $\bigcup_{m,l} L_{m,l}$ is the set $L^+(f)$ of all points having positive Lyapunov exponent. Using the Lebesgue density theorem and the absolute continuity of μ we take off every set $L_{m,l}$ a set of zero μ-measure so that all points of the resultant set $L'_{m,l}$ are density points. In particular, every neighbourhood in $L'_{m,l}$ of each of its points has positive Lebesgue measure. We can assume that $L'_{m,l} \subset \operatorname{supp} \mu$ for every m and l.

Fix arbitrarily m and l. Let k be an integer such that $(1 + 1/m)^k > 6$ and find a small $\delta > 0$ so that:

- the distance between consecutive points of P is greater than δ;
- if J is a δ-singular interval and $i < k$, then either $f^i(J)$ is just one point or $f^i(J) \cap P = \emptyset$.

We prove that all points from $L'_{m,l}$ have the required property in the lemma for this δ. Since $\mu(\bigcup_{m,l} L'_{m,l}) = \mu(L^+(f))$, the lemma follows.

Assume to the contrary that there is a point $x \in L'_{m,l}$ and an interval U neighbouring x, $\lambda(U) < \delta$, with the property that if J is a substantial subinterval of U, $f^r(J) \cap P \neq \emptyset$ for some $r \geq 1$, and $f^i(J) \cap P = \emptyset$ for every $0 \leq i < r$, then $f^r(J) \subset (a-\delta, a+\delta)$ for some $a \in P$. In particular, $f^r(J)$ intersects at most three elements from the partition $\mathcal{P}$ consisting of the points of P and the intervals having them as their endpoints.

Let $A = L'_{m,l} \cap U$ and recall that $\lambda(A) > 0$. Clearly, the number of sets from the partition $\mathcal{P}^n = \bigvee_{i=0}^{n-1} f^{-i}(\mathcal{P})$ containing points from A is less than $3^{1+n/k}$, which means that for every n there is a set $A_n \subset A$ with $\lambda(A_n) > \lambda(A)/3^{1+n/k}$ that is contained in some $P_n \in \mathcal{P}^n$. Observe that P_n is a nondegenerate interval and that f^n is continuous and monotone on P_n. Further, $A_n \subset L_{m,l}$ implies that no point of A_n belongs to an interval of constancy of f^n. In particular, f^n is one-to-one on A_n. Therefore we have

$$\lambda(f^n(A_n)) > (1+1/m)^n \lambda(A_n) > \lambda(A)2^{n/k}/3$$

for every $n \geq l$. If n is large enough, we arrive at the contradiction $\lambda(f^n(A_n)) > 1$. □

Let P be a singular set. We introduce the set of symbols $P_\pm = \{a_-, a_+ : a \in P\} \setminus \{0_-, 1_+\}$. If $\delta > 0$ is less than the distance between every pair of consecutive points of P, then we call each of the intervals $P_\delta(a_-) := (a-\delta, a)$ and $P_\delta(a_+) := (a, a+\delta)$ a *δ-singular interval (for P and f)*, or just a *singular interval* if no emphasis on δ is needed.

We say that $b \in P_\pm$ is *expanding* if there is a number δ such that, for every $\epsilon > 0$, there is a positive integer n such that the length of some component of $f^n(P_\epsilon(b))$ is greater than δ.

Finally, an interval $J \subset I$ is said to be a *homterval* for f if $f|_{f^n(J)}$ is continuous and monotone for every $n \geq 0$.

Lemma 5.2. *Let $f : I \to I$ be a piecewise monotone map having no homtervals and let P be a singular set for f.*

- *(i) If every $b \in P_\pm$ is expanding, then there is $\delta > 0$ such that if J is a subinterval of I and n is large enough then the length of some component of $f^n(J)$ is greater than δ.*
- *(ii) If the set $N \subset P_\pm$ of non-expanding symbols is non-empty, then for every $\epsilon > 0$ and every $b \in N$ there are numbers $0 < \nu_{b,\epsilon} \leq \epsilon$ with the following property: if $b \in N$ and $k = k(b,\epsilon)$ is the first positive integer such that $f^k(P_{\nu_{b,\epsilon}}(b))$ intersects P, then $\lambda(f^i(P_{\nu_{b,\epsilon}}(b))) \leq \epsilon$ for every $0 \leq i < k$ and $f^k(P_{\nu_{b,\epsilon}}(b)) \subset N \cup \bigcup_{b' \in N} P_{\nu_{b',\epsilon}}(b')$.*

Proof. We prove (i). Since every $b \in P_\pm$ is expanding there is a number $\delta' > 0$ with the property that, for every $\epsilon > 0$ and every $b \in P_\pm$, there is a positive integer n (depending on ϵ and b) such that some component of $f^n(P_\epsilon(b))$ has length less than δ'. Let $\mathcal{Q}$ be a partition of I into intervals of lengths smaller than $\delta'/2$. The absence of homtervals implies that there is a finite family $\mathcal{R}$ of singular subintervals of I with the following property: if $Q \in \mathcal{Q}$, then $f^n(Q) \supset R$ for some $R \in \mathcal{R}$ and $n = n(Q) \geq 1$. On the other hand if $R \in \mathcal{R}$, then, since all $b \in P_\pm$ are expanding, $f^m(R) \supset Q$ for some $Q \in \mathcal{Q}$ and $m = m(R) \geq 1$.

Let $\delta > 0$ be such that all sets $f^i(Q)$, $0 \leq i < n(Q)$, and $f^j(R)$, $0 \leq j < m(R)$, have some component of length greater than δ. If J is a subinterval of I then there are some $k \geq 0$, $b \in P_\pm$ and $\epsilon > 0$ such that $f^k(J) \supset P_\epsilon(b)$, and hence some $l \geq k$ and $Q \in \mathcal{Q}$ such that $f^l(J) \supset Q$. Thus, if $n \geq l$, some component of $f^n(J)$ has length greater than δ. This proves (i).

Now we prove (ii). We assume $\epsilon > 0$ to be smaller than the minimum distance between consecutive points of P. We can further assume that if all components of all sets $f^n(P_\nu(b))$, $n \geq 0$, have length at most ϵ for some number $\nu > 0$ and some symbol $b \in P_\pm$, then $b \in N$. For every $b \in N$ let $\nu_{b,\epsilon}$ be the maximum number ν with this property. We claim that the numbers $\nu_{b,\epsilon}$ do the work.

Suppose not to find a symbol $b \in N$ such that $f^k(P_{\nu_{b,\epsilon}}(b)) \not\subset N \cup \bigcup_{b' \in N} P_{\nu_{b',\epsilon}}(b')$ for the first positive integer k such that $f^k(P_{\nu_{b,\epsilon}}(b))$ intersects P (such a number k exists because of the absence of homtervals for f). Let $\nu < \nu_{b,\epsilon}$ be close enough to $\nu_{b,\epsilon}$ so that $f^i(P_\nu(b)) \cap P = \emptyset$ for every $0 \leq i < k$ and $f^k(P_\nu(b)) \not\subset N \cup \bigcup_{b' \in N} P_{\nu_{b',\epsilon}}(b')$. Then there is a subinterval J of $P_\nu(b)$ satisfying $f^k(J) = P_\delta(c)$ such that either $c \in P_\pm \setminus N$ or $\delta > \nu(c,\epsilon)$. Since all components of all sets $f^n(P_\nu(b))$ have length at most ϵ, the same is true for the components of the sets $f^n(P_\delta(c))$. Hence $c \in N$ and $\delta \leq \nu(c,\epsilon)$, a contradiction. □

Remark 5.3. If the numbers $\nu(b,\epsilon)$ are chosen as in the proof of Lemma 5.2(ii), then none of the intervals $f^i(P_{\nu_{b,\epsilon}}(b))$, $0 \leq i < k(b,\epsilon)$, can be a singular interval $P_\delta(c)$ for some expanding c. Consequently, if we denote $A_\epsilon = \bigcup_{b \in N} \bigcup_{i=0}^{k(b,\epsilon)-1} f^i(P_{\nu_{b,\epsilon}}(b))$, then there is a number $d = d(\epsilon)$ such that $A_\epsilon \cap P_d(c) = \emptyset$ for every expanding symbol c.

Proof of Theorem B. The first statement of Theorem B follows from Lemma 5.1.

We prove now the second statement of Theorem B. The hypotheses $\mu(L^+(f)) = 1$ and $\operatorname{supp}\mu = I$ prevent f to have homtervals. Indeed, if J is a homterval then $\mu(J) > 0$ (because $\mu(O) > 0$ for every open set O). Since $\mu(L^+(f)) = 1$, $L_{m,l} \cap J$ has positive μ-measure (and therefore, by the absolute continuity of μ, positive Lebesgue measure) for some of the sets

$$L_{m,l} = \{x \in I : |(f^n)'(x)| > (1 + 1/m)^n \text{ for every } n \geq l\}.$$

Since f^n is one-to-one on $L_{m,l} \cap J$ for every n, we get $\lambda(f^n(L_{m,l} \cap J)) > (1 + 1/m)^n \lambda(L_{m,l} \cap J)$ for every $n \geq l$, a contradiction.

Now we assume that one of the conditions (a), (b) or (c) holds. Let P be a singular set for f (if (a) holds then we choose it with the additional property that the restriction of f to every pair of consecutive points of P is a C^2-diffeomorphism; if (c) holds then P is the set E of discontinuities and local extrema of f). In view of Lemma 5.2(i) it suffices to show that every $b \in P_\pm$ is expanding. Suppose not and denote by N the set of non-expanding symbols.

Notice that, because of the absolute continuity of μ, the set T of points such that $f^n(x) \notin P$ for all integers $n \geq 0$ has full μ-measure. Let

$$A_\epsilon = \bigcup_{b \in N} \bigcup_{i=0}^{k(b,\epsilon)-1} f^i(P_{\nu_{b,\epsilon}}(b))$$

for every $\epsilon > 0$ (Lemma 5.2(ii)). Then A_ϵ is a finite union of open intervals containing no singular points, it is "almost" invariant by f (in the sense that $f(T \cap A_\epsilon) \subset T \cap A_\epsilon$), and it is at a "positive distance" of all expanding symbols (Remark 5.3).

We must consider separately the cases when (a), (b) or (c) holds. If (c) holds, then we have $P = E$ and $f(A_\epsilon) \subset A_\epsilon$ whenever ϵ is small enough. This implies the existence of homtervals for f, a contradiction.

Henceforth we can assume that (a) or (b) holds. Fix $\epsilon_0 > 0$ and notice that $A_\epsilon \subset A_{\epsilon_0}$ for all small numbers $\epsilon > 0$. We claim that if ϵ is one of such numbers, then

$$\text{for } \mu\text{-a.e. } x \in A_{\epsilon_0} \text{ there is an } n(x) \geq 0 \text{ such that } f^{n(x)}(x) \in A_\epsilon. \tag{5}$$

Assume that (a) holds. Recall that if $x \in T \cap A_{\epsilon_0}$, then none of its iterates is "near" to an expanding singular symbol. Then (5) follows from a result by Mañé [26] according to which if f is piecewise C^2-diffeomorphic with respect to P and it has no homtervals then, for every $\delta > 0$, the set of points x such that $\mathrm{dist}(f^n(x), P) \geq \delta$ for every $n \geq 0$ has zero Lebesgue measure and hence zero μ-measure. If (b) holds, then we can prove an even stronger fact: $A_{\epsilon_0} \setminus A_\epsilon$ is finite. In the opposite case there would exist an open interval $J \subset A_{\epsilon_0} \setminus A_\epsilon$ (for both A_{ϵ_0} and A_ϵ are finite union of intervals). Taking if necessary a smaller interval we can assume that J is homeomorphically mapped by some f^l onto a singular interval $K \subset A_\epsilon$. Recall that the set of recurrent points of f has full μ-measure because μ is invariant for f. In particular, if A is the set of recurrent points from $T \cap J$, then $\mu(J) > 0$ implies $\mu(A) > 0$, hence $A \neq \emptyset$. Since $f^n(A) \subset A_\epsilon$ for every $n \geq l$, we see that A cannot contain any recurrent point, a contradiction.

Thus if (a) or (b) holds, then (5) holds as well. Since $\mu(L^+(f)) = 1$, we can use Lemma 5.1 to find $A \subset T \cap A_{\epsilon_0}$ with $\mu(A) > 0$ and a number $\delta > 0$ such that if $x \in A$ and U is a neighbourhood of x, then some component of some iterate of U has length greater than δ. Take $\epsilon < \delta$ and use (5) to find an $x \in A$ such that $f^m(x) \in A_\epsilon$ for some m. Notice that f is continuous at all points of the orbit of x (because $x \in T$), so we can find a small neighbourhood U of x and a number l such that f^l maps homeomorphically U into some interval $P_\nu(b, \epsilon)(b)$ and $\lambda(f^i(U)) \leq \epsilon$ for every $0 \leq i < l$. Now we apply Lemma 5.2(ii) to get that, in fact, all components of all iterates of U have length at most ϵ. This is a contradiction. □

6. **Proof of Theorem C.** In this section we construct a C^1-map $g : I \to I$ with exactly one critical point satisfying $\lambda(L^+(g)) = 1$ but not having full sensitivity to initial conditions.

Let $(\kappa_n)_{n=0}^\infty$ be a sequence of positive numbers such that:

- $\kappa_1 > 2\kappa_0^2/(1 - \kappa_0)$;
- $\sum_n \kappa_n < 1$;
- the sequence (κ_n/κ_{n+1}) decreases to 1;

for instance one could take $\kappa_n = 1/(n+3)^2$. Let $x_0 = 1/15$, $x_1 = 4/15$, $x_2 = 7/15$, $x_3 = 2/3$, $x_4 = 13/15$, and $x_5 = 29/30$. Write $I_{0,1} = [x_3, x_4]$, $J_{0,1} = (x_2, x_3)$, $\alpha_0 = \lambda(I_{0,1}) = 1/5$, $\beta_0 = \lambda(J_{0,1}) = 1/5$, and define, for every positive integer n, numbers α_n and β_n, compact intervals $(I_{n,i})_{i=1}^{2^n}$, and open intervals $(J_{n,j})_{j=1}^{2^{n-1}}$ satisfying the following properties:

- $\lambda(I_{n,i}) = \alpha_n$, $\lambda(J_{n,j}) = \beta_n$;
- every interval $I_{n-1,i}$ is the disjoint union of the consecutive intervals (from left to right) $I_{n,2i-1}$, $J_{n,i}$, and $I_{n,2i}$;
- $\beta_n = \kappa_{n-1}\alpha_{n-1}$.

Observe that

$$\frac{\alpha_n}{\alpha_{n+1}} = \frac{2}{1 - \kappa_n}$$

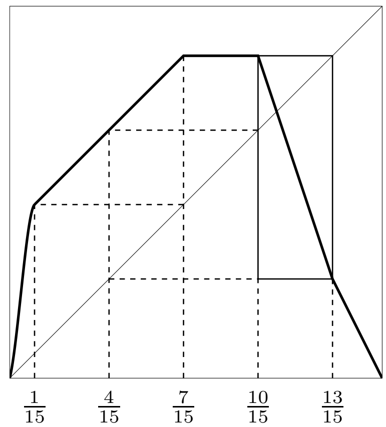

FIGURE 1. The graph of the map f_0.

for every $n \geq 0$, and

$$\frac{\beta_n}{\beta_{n+1}} = \frac{2}{1-\kappa_{n-1}}\frac{\kappa_{n-1}}{\kappa_n}$$

for every $n \geq 1$. Also,

$$\frac{\beta_0}{\beta_1} = \frac{1}{\kappa_0} > \frac{2\kappa_0}{(1-\kappa_0)\kappa_1} = \frac{\beta_1}{\beta_2}$$

by hypothesis. Hence

$$\text{the sequences } (\alpha_n/\alpha_{n+1}) \text{ and } (\beta_n/\beta_{n+1}) \text{ decrease to 2.} \tag{6}$$

Thus, if we define $B = \bigcap_{n=0}^{\infty}\bigcup_{i=1}^{2^n} I_{n,i}$, then (6) implies that B is a Cantor type set. Moreover, notice that

$$\lambda\left(\bigcup_{i=1}^{2^n} I_{n,i}\right) = 2^n\alpha_n = 2^{n-1}(1-\kappa_{n-1})\alpha_{n-1} = \cdots = \frac{1}{5}\prod_{l=0}^{n-1} 1-\kappa_l.$$

Since $\sum_n \kappa_n < 1$, we get $\lambda(B) > 0$.

We construct our map g via a double iterative process. First we define inductively a sequence (f_n) of continuous interval maps. Let $\varphi : [0, x_0] \to [0, x_2]$ be an increasing diffeomorphism with negative Schwarzian derivative satisfying $\varphi'(0) = 4$ and $\varphi'(x_0) = 1$. Notice that such a map exists by Lemma 3.15. We define $f_0 : I \to I$ by

$$f_0(x) = \begin{cases} \varphi(x) & \text{if } x \in [0, x_0], \\ x + 6/15 & \text{if } x \in [x_0, x_2], \\ 13/15 & \text{if } x \in [x_2, x_3], \\ 43/15 - 3x & \text{if } x \in [x_3, x_4], \\ 2 - 2x & \text{if } x \in [x_4, 1]. \end{cases}$$

See Figure 1.

Assume that f_n is already defined. Then we define f_{n+1} so that it equals f_n outside the intervals $I_{n,i}$. Further, if $1 \leq j \leq 2^n$, then the interval $I_{n+1,j}$ is mapped affinely by f_{n+1} onto the interval $I_{n,2^n-j+1}$, and if $2^n + 1 \leq j \leq 2^{n+1}$, then then interval $I_{n+1,j}$ is mapped affinely by f_{n+1} into the interval $[x_1, x_2]$, and by f_{n+1}^2 onto the interval $I_{n,2^{n+1}-j+1}$ for every . In particular, observe that every interval

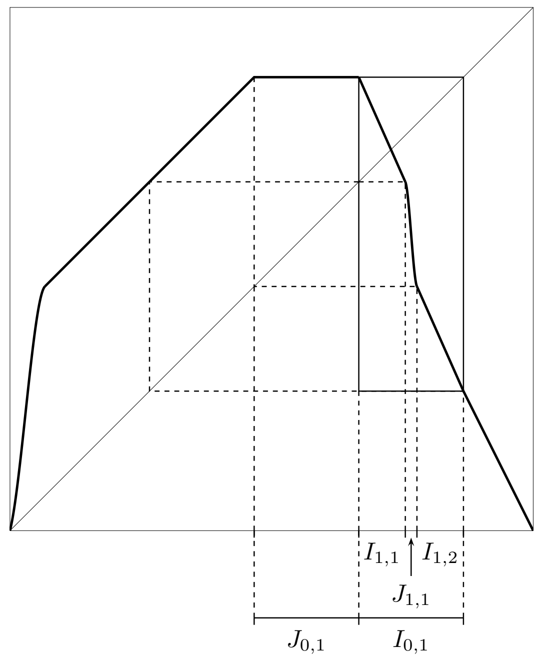

FIGURE 2. The graph of the map f_1.

$I_{n+1,j}$ (which has length α_{n+1}) is mapped onto an interval of length α_n. Finally we use Lemma 3.15 to define f_{n+1} on every interval $J_{n+1,j}$, so that it has negative Schwarzian derivative on the whole interval $J_{n+1,j}$ and derivative -2 at its endpoints. We emphasize that every interval $J_{n+1,j}$ (which has length β_{n+1}) is either mapped by f_{n+1} onto some interval $J_{n,k}$, or into $[x_1, x_2]$ and then onto some interval $J_{n,k}$. Thus, in any case, the corresponding interval $f_{n+1}(J_{n+1,j})$ has length β_n, the ratio $\lambda(f_{n+1}(J_{n+1,j}))/\lambda(J_{n+1,j})$ is greater than 2 (recall (6)), and we can safely use Lemma 3.15. See Figure 2.

Clearly, the sequence (f_n) converges uniformly to a map f that equals f_0 outside $[x_3, x_4]$ and maps homeomorphically $[x_3, x_4]$ onto $[x_1, x_4]$. Moreover, the ratios $\lambda(f(I_{n,i}))/\lambda(I_{n,i}) = \alpha_{n+1}/\alpha_n$ and $\lambda(f(J_{n,j}))/\lambda(J_{n,j}) = \beta_{n+1}/\beta_n$ tend to 2 as $n \to \infty$ by (6), which guarantees that $f'(x) = -2$ for every $x \in B \setminus \{x_3, x_4\}$, and the same is true for the corresponding one-sided derivatives $f'_r(x_3)$ and $f'_l(x_4)$. (Indeed we have $f'(x_4) = -2$ because of the way f was defined at $[x_4, 1]$.) Now it is important to realize that f', when restricted to the intervals $J_{n,j}$, goes uniformly to -2 as $n \to \infty$, because we used Lemma 3.15 to define our maps f_n there, hence property Lemma 3.15(iii) applies. We conclude that the restriction of f to $I \setminus [x_2, x_3]$ is a C^1-map.

Finally, we get our desired map g by taking f as the starting point of a new iterative process, based on the $*$-operation we next introduce. Let $h : I \to I$ be a continuous map satisfying $h(0) = h(1) = 0$. Then the map $f * h : I \to$ is defined by

$$(f * h)(x) = \begin{cases} 13/15 + g(10/3 - 5x)/10 & \text{if } x \in [x_2, x_3], \\ f(x) & \text{otherwise.} \end{cases}$$

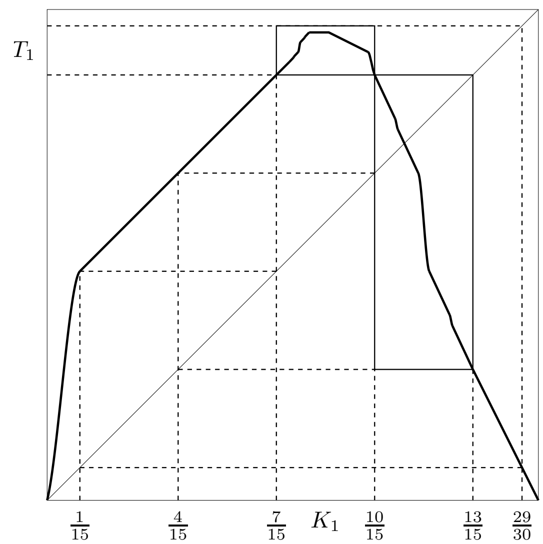

FIGURE 3. The graph of the map $g_2 = f * f$.

Thus, the graph of $f * h$ consists of the symmetrical rescaled graph of g inside the rectangle $[x_2, x_3] \times [x_4, x_5]$ and the graph of f outside it. Notice that $(f * h)^3$ maps $[x_2, x_3]$ into itself and its restriction to this interval is topologically conjugated to h via a decreasing affine map. More precisely, if $\rho : [x_2, x_3] \to [0, 1]$ is defined by $\rho(x) = (x_3 - x)/(x_3 - x_2)$, then $(f * h)^3|_{[x_2,x_3]} = \rho^{-1} \circ h \circ \rho$.

Let $\sigma : [x_4, x_5] \to [0, 1]$ be given by $\sigma(y) = (y - x_4)/(x_5 - x_4)$. Let $K_1 = [x_2, x_3]$, $T_1 = [x_4, x_5]$, and $g_1 = f$, and define inductively $K_m = \rho^{-1}(K_{m-1})$, $T_m = \sigma^{-1}(T_{m-1})$, and $g_m = f * g_{m-1}$ for every $m > 1$. It is easy to check that (K_m) and (T_m) are decreasing sequences of intervals, the lengths of K_m and T_m being, respectively, one-fifth and one-tenth of those of K_{m-1} and T_{m-1}. Moreover, every g_m equals g_{m-1} outside K_{m-1} and consists of the rescaled graph of g (if m is odd) or the symmetrical rescaled graph of g (if m is even) inside the rectangle $K_m \times T_m$. See Figure 3.

Thus (g_m) converges to a map g having exactly one turning point c, the intersection point of all intervals K_m, with $g(c)$ being the intersection point of all intervals T_m. Notice that the derivative of g is well defined at the points connecting the consecutive "pieces" the function g is made of. The reason is that the length of K_1 is twice that of T_1, while $f'(0) = 4 = -2f'_r(x_3)$ and $f'(1) = 2 = -2f'_l(x_2)$. This, together with the fact that the ratios $\lambda(T_m)/\lambda(K_m)$ go to zero as $m \to \infty$, implies that g is a C^1-map having c as its only critical point. We claim that it is the map we are looking for.

The map g has a key feature: $g * g = g$. As a consequence, every g^{3^m} maps K_m into itself and its restriction to K_m is affinely conjugated to g itself. Furthermore, it is plain to see that the lengths of the intervals $f^i(K_m)$, $0 \leq i < 3^m$, decrease uniformly to zero. Then c cannot be sensitive and g cannot have full sensitivity.

Now we must show that $\lambda(L^+(g)) = 1$. Let $C = B \cap I_{1,1}$, $R = B \cap I_{1,2}$, and $L = g(R)$. Recall that g' equals 1 in L and equals -2 both in C and R, and notice that $g(L) = g(C) = C \cup R$. Let μ be the probability measure defined by

$$\mu(A) = \frac{\lambda(A \cap L)}{3\lambda(L)} + \frac{\lambda(A \cap C)}{3\lambda(C)} + \frac{\lambda(A \cap R)}{3\lambda(R)}$$

for every Borel set A. We claim that μ is invariant for g (in fact it is an acip, its density being $\rho = 1_{L\cup C\cup R}/\lambda(L \cup C \cup R)$, but this is not important here). If suffices to check that $\mu(g^{-1}(A)) = \mu(A)$ for every set A that is a subset either of L, C, or R. For instance, assume that $A \subset C$ (the cases $A \subset L$ and $A \subset R$ can be dealt with in similar fashion). Then there are sets $A_L \subset L$ and $A_C \subset C$ such that $g(A_L) = g(A_C) = A$ and $A_L \cup A_C = g^{-1}(A)$. Thus,

$$\begin{aligned} \mu(g^{-1}(A)) &= \mu(A_L) + \mu(A_C) \\ &= \frac{\lambda(A_L)}{3\lambda(L)} + \frac{\lambda(A_C)}{3\lambda(C)} = \frac{\lambda(A)}{6\lambda(C)} + \frac{\lambda(A)/2}{3\lambda(C)} = \frac{\lambda(A)}{3\lambda(C)} = \mu(A). \end{aligned}$$

We have proved that μ is an acip for g. Since μ and λ have the same zero measure sets in $D_0 := L \cup C \cup R$ and $\log|f'|$ is μ-integrable, Birkhoff's ergodic theorem (Theorem 3.4) implies that $\Lambda_g(x)$ is well defined Λ-a.e. $x \in D_0$. We have $g(D_0) = D_0$ and $g'(x) = 1$ or $g'(x) = -2$ for every $x \in D_0$. Moreover, $g'(x) = 1$ implies $g'(g(x)) = -2$ for every such x. Then we get $\Lambda_g(x) > 0$ for λ-a.e. $x \in D_0$.

Recall that the restriction of g^{3^m} to K_m is conjugated to g via an affine map $\rho_m : K_m \to I$ (so ρ_1 is the previously defined map ρ). Write $D_m = \rho_m^{-1}(D_0)$ and fix m. From our previous reasoning we know that

$$\lim_{l\to\infty} \frac{\log|(g^{3^m l})'(x)|}{3^m l} > 0 \quad \text{for } \lambda\text{-a.e. } x \in D_m.$$

Observe that there are positive numbers r_m depending only on m with $1/r_m < |g'(g^k(x))| < r_m$ for every $x \in D_m$ and every $k \geq 0$. Consequently if $x \in D_m$, then

$$\frac{|(g^{3^m l})'(x)|}{r_m^{-3^m}} < |(g^k)'(x)| < |(g^{3^m l})'(x)| r_m^{3^m}$$

and

$$\frac{\log|(g^{3^m l})'(x)|}{3^m(l+1)} - \frac{\log r_m}{l+1} < \frac{\log|(g^k)'(x)|}{k} < \frac{\log|(g^{3^m l})'(x)|}{3^m l} + \frac{\log r_m}{l}$$

whenever $3^m l \leq k < 3^m(l+1)$. Therefore,

$$\lim_{k\to\infty} \frac{\log|(g^k)'(x)|}{k} = \lim_{l\to\infty} \frac{\log|(g^{3^m l})'(x)|}{3^m l} > 0 \quad \text{for } \lambda\text{-a.e. } x \in D_m.$$

Since g is a C^1-map with a single singular point, it is non-singular. Therefore, to finish the proof it suffices to show that for λ-a.e. $x \in I$ there are some nonnegative numbers k and m satisfying $g^k(x) \in D_m$. In other words, if we define the sets $O_m = \bigcap_{l=0}^{\infty} g^{-l}\left(I \setminus \bigcup_{i=0}^{m} D_i\right)$, then we must prove that $\lambda(O_m) \to 0$ as $m \to \infty$. We do this by showing that if U is a component of $O_m \setminus \{0,1\}$ for some m, then there is a number $\epsilon > 0$ not depending on m such that $\lambda((I \setminus O_{m+1}) \cap U)/\lambda(U) > \epsilon$.

Let U be such interval. Because of the definition of g, there is an iterate of g diffeomorphically mapping the closure of U onto K_m; moreover, either this diffeomorphism is affine, or it has negative Schwarzian derivative. Observe that K_m is a 2/15-scaled neighbourhood of an interval containing D_{m+1} and that the number $\lambda(D_{m+1})/\lambda(K_m)$ does not depend on m. Then the existence of the required number ϵ follows from Lemma 3.14.

7. **Positive Lyapunov exponents and positive entropy are equivalent: The Rohlin-Ledrappier formula.** This section is organized as follows. First we just assume invariance for μ to get the Margulis-Ruelle inequality; then we add absolutely continuity and prove a number of lemmas from which the Rohlin-Ledrappier formula will immediately follow.

We need this version of the inequality (recall that $\chi_f = \Lambda_f^+ = \max\{\Lambda_f, 0\}$):

Theorem 7.1 (the Margulis-Ruelle inequality)**.** *Let $f : I \to I$ be a piecewise Lipschitz map, let μ be an invariant probability measure for f, and assume that f is differentiable μ-a.e. Then $h_\mu(f) \leq \int_I \chi_f \, d\mu < \infty$.*

We emphasize that every piecewise Lipschitz map is piecewise absolutely continuous, hence it is differentiable λ-a.e. Thus the μ-a.e differentiability condition is automatically satisfied when μ is an acip.

It is worth noticing that Ledrappier misquotes the Margulis-Ruelle inequality in [23] by saying that if f is C^1 and μ is invariant for f, then $h_\mu(f) \leq \max\{0, \int_I \log|f'| \, d\mu\}$. Curiously enough Mañé, when referring to Ledrappier's result in [27, p. 230], misquotes the inequality in a different way: he claims that if $\log|f'|$ is μ-integrable, then $h_\mu(f) \leq \int_I \log|f'| \, d\mu$. Notice that none of these statements works. Let f be the C^1-map defined by

$$f(x) = \begin{cases} x/2 + 32x^7 & \text{if } 0 \leq x \leq 1/2, \\ 1/2 + 8(x - 1/2)(1 - x) & \text{if } 1/2 \leq x \leq 1 \end{cases}$$

(see Figure 4). Then the orbits of all points from $[0, 1/2)$ tend to 0, while in $[1/2, 1]$ f is just the resized full parabola $g(x) = 4x(1 - x)$. It is a standard fact that g admits exactly one acip μ' ([28, Theorem 1.5, p. 349]). Moreover, the Lebesgue measure λ is invariant for the tent map $T(y) = 1 - |2y - 1|$ and the conjugacy $\varphi(y) = \sin^2(\pi y/2)$ preserves the measures (see, e.g., [28, pp. 107 and 352]). In particular $h_{\mu'}(g) = h_\lambda(T) = \log 2$ (the second equality follows for instance from the Rohlin-Ledrappier formula). Hence, if we define the invariant measure μ by $\mu(\{0\}) = 1/2$ and $\mu(A) = \mu'(2A - 1)/2$ whenever $A \subset [1/2, 1]$ is a Borel set (here $2A - 1 = \{2x - 1 : x \in A\}$), then we get $h_\mu(f) = \log 2/2$ (Proposition 3.3(ii)). On the other hand, $\int_I \log|f'| \, d\mu = 0$ because

$$\int_{[0,1/2)} \log|f'| \, d\mu = \int_{\{0\}} \log|f'| \, d\mu = -\log 2/2$$

and, using the Rohlin-Ledrappier formula,

$$\int_{[1/2,1]} \log|f'| \, d\mu = \frac{1}{2} \int_I \log|g'| \, d\mu' = \log 2/2.$$

Proof of Theorem 7.1. For completeness we provide a detailed proof. We follow closely [27, Section IV.12, pp. 281–285].

To begin with notice that $\log|f'|$ has μ-integral because of the hypotheses on f and μ. Hence χ_f is well defined μ-a.e. (apply Theorem 3.4 to $u = \log|f'|$). Indeed, if $L > 1$ is a valid Lipschitz constant for f on each of its intervals of continuity, then $0 \leq \chi_f \leq \log L$ μ-a.e, so χ_f is μ-integrable.

Let $(\mathcal{P}_m)_{m=1}^\infty$ be a increasing sequence of interval partitions of I such that

$$\text{all intervals from } \mathcal{P}_m \text{ have the same length,} \tag{7}$$

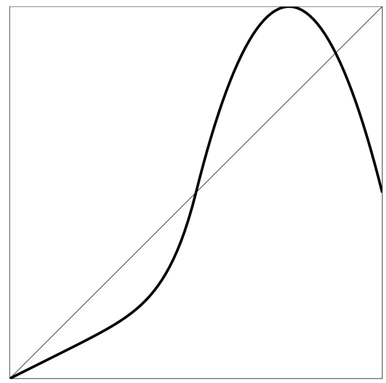

FIGURE 4. The graph of the map f.

and

$$\operatorname{diam} \mathcal{P}_m \to 0 \text{ as } m \to \infty. \tag{8}$$

We denote by $\mathcal{P}_m$ the σ-algebra generated by $\mathcal{P}_m$. If $x \in I$, then $P_m(x)$ denotes the interval from $\mathcal{P}_m$ containing x. If $g : I \to I$ is a map, then we write

$$\nu_{g,m}(x) = \operatorname{Card}\{P \in \mathcal{P}_m : g(P_m(x)) \cap P \neq \emptyset\}$$

and

$$\nu_g(x) = \limsup_{m\to\infty} \nu_{g,m}(x).$$

Observe that if d is the number of discontinuity points of f and P is an interval from some partition $\mathcal{P}_m$, then $f(P)$ intersects at most $C = (d+1)(L+2)+d$ intervals from $\mathcal{P}_m$ (we use (7)). From this

$$\nu_{f^n,m}(x) \le C^n \text{ for every } x \in I \text{ and every } n, m \ge 1,$$

so

$$\frac{\log \nu_{f^n}}{n}, \frac{\log \nu_{f^n,m}}{n} \le \log C \text{ for every } n, m \ge 1. \tag{9}$$

Also, notice that if g is differentiable at a point x and $[y]$ denotes the integer part of a number y, then $[|g'(x)|] \le \nu_g(x) \le [|g'(x)|] + 2$ because of (8), which implies

$$\lim_{n\to\infty} \frac{\log \nu_{f^n}}{n} = \chi_f \quad \mu\text{-a.e.} \tag{10}$$

Next we claim that

$$h_\mu(f^n, \mathcal{P}_m) \le \int_I \log \nu_{f^n,m}\, d\mu \quad \text{for every } n, m \ge 1. \tag{11}$$

To prove (11) we write $g = f^n$ and recall that

$$\begin{aligned} h_\mu(g, \mathcal{P}_m) &= \lim_{r\to\infty} \frac{1}{r} H_\mu\left(\bigvee_{l=0}^{r} g^{-l}(\mathcal{P}_m)\right) \\ &= \lim_{r\to\infty} \frac{1}{r}\left[H_\mu(\mathcal{P}_m) + \sum_{s=1}^{r} H_\mu\left(g^{-s}(\mathcal{P}_m) \middle| \bigvee_{n=0}^{s-1} g^{-n}(\mathcal{P}_m)\right)\right] \\ &= \lim_{r\to\infty} \frac{1}{r} \sum_{s=1}^{r} H_\mu\left(g^{-s}(\mathcal{P}_m) \middle| \bigvee_{n=0}^{s-1} g^{-n}(\mathcal{P}_m)\right); \end{aligned}$$

we have also used Proposition 3.1(i). Thus, if we define

$$\delta_s(x) = -\sum_{P_s \in g^{-s}(\mathcal{P}_m)} \frac{\mu(Q \cap P_s)}{\mu(Q)} \log \frac{\mu(Q \cap P_s)}{\mu(Q)},$$

for every $s \geq 1$ and $x \in I$, where $Q = Q(x)$ is given by the property that $x \in Q = P_0 \cap P_1 \cap \cdots \cap P_{s-1}$ with $P_l \in g^{-l}(\mathcal{P}_m)$ for every $0 \leq l < s$, then we get

$$h_\mu(g, \mathcal{P}_m) = \lim_{r \to \infty} \frac{1}{r} \int_I \sum_{s=1}^{r} \delta_s \, d\mu.$$

Moreover, observe that the number of nonzero summands in the definition of $\delta_s(x)$ does not exceed $\nu_{g,m}(g^{s-1}(x))$. Hence

$$h_\mu(g, \mathcal{P}_m) \leq \lim_{r \to \infty} \int_I \frac{1}{r} \sum_{s=1}^{r} \log(\nu_{g,m} \circ g^{s-1}) \, d\mu = \int_I \lim_{r \to \infty} \frac{1}{r} \sum_{s=1}^{r} \log(\nu_{g,m} \circ g^{s-1}) \, d\mu$$

by (1), (11) and the dominated convergence theorem. Now (11) follows from Birkhoff's ergodic theorem.

We have

$$h_\mu(f) = \frac{h_\mu(f^n)}{n} \quad \text{for every } n \geq 1$$

by Proposition 3.3(i). Also, because of (8), Proposition 3.3(iv) implies

$$h_\mu(f^n) = \lim_{m \to \infty} h_\mu(f^n, \mathcal{P}_m) \quad \text{for every } n, m \geq 1.$$

Furthermore, using (9) and the dominated convergence theorem we get

$$\limsup_{m \to \infty} \int_I \frac{\log \nu_{f^n,m}}{n} \, d\mu \leq \int_I \limsup_{m \to \infty} \frac{\log \nu_{f^n,m}}{n} \, d\mu = \int_I \frac{\log \nu_{f^n}}{n} \, d\mu.$$

Hence (11) implies

$$h_\mu(f) \leq \int_I \frac{\log \nu_{f^n}}{n} \, d\mu \quad \text{for every } n \geq 1.$$

Since

$$\lim_{n \to \infty} \int_I \frac{\log \nu_{f^n}}{n} \, d\mu = \int_I \lim_{n \to \infty} \frac{\log \nu_{f^n}}{n} \, d\mu = \int_I \chi_f \, d\mu$$

by (9) and (10), the lemma follows. □

In the next lemmas we adapt the original arguments by Rohlin and Ledrappier to complete the proof of Theorem D. In doing this, we found [10, Lemma 1.4] particularly enlightening. Until the end of the proof of Lemma 7.4, the piecewise Lipschitz map $f : I \to I$ and its acip μ will remain fixed.

Lemma 7.2. *Let $g = \log|f'| - \log \rho + \log \rho \circ f$, with ρ being the density of the measure μ. Then $g \geq 0$ μ-a.e. Moreover, $\int_I g \, d\mu \leq h_\mu(f)$.*

Remark 7.3. Notice that if we write $D = \{x : \rho(x) = 0\}$ and $C = \{y : \rho(f(y)) = 0\}$, then we have $C = f^{-1}(D)$. Thus

$$\mu(C) = \mu(D) = \int_D \rho \, d\lambda = 0$$

and g is well defined μ-a.e.

Proof of Lemma 7.2. Fix $\epsilon > 0$ and $r \in \mathbb{N}$. With the notation of Theorem 3.13, we consider the partition $\mathcal{P} = \{K_1, \dots, K_r, N\}$, $N = I \setminus \bigcup_{j=1}^{r} K_j$. If the number r is large enough, then $\lambda(f(N))$ is very small due to Theorem 3.13(i). Since μ is absolutely continuous, $\mu(f(N))$ is also very small; in particular, we can assume $\mu(f(N)) < \epsilon$. Since μ is invariant, the μ-measure of $K = f^{-1}(I \setminus f(N))$ is greater than $1 - \epsilon$. We show that

$$g(x) \geq 0 \quad \text{for } \mu\text{-a.e } x \in K \tag{12}$$

and

$$\int_K g \, d\mu \leq h_\mu(f). \tag{13}$$

Since ϵ is arbitrary, this implies the lemma.

Find partitions $(\mathcal{P}_m)_{m=1}^{\infty}$ with the properties

$$\{f(N), I \setminus f(N)\} \leq \mathcal{P}_m \quad \text{for every } m \geq 1, \tag{14}$$

$$\bigvee_{n=0}^{m-1} f^{-n}(\mathcal{P}) \leq \mathcal{P}_m \quad \text{for every } m \geq 1, \tag{15}$$

and

$$\operatorname{diam} \mathcal{P}_m \to 0 \text{ as } m \to \infty, \tag{16}$$

and define $\mathcal{F}_m = f^{-1}(\mathcal{P}_m)$ and $\mathcal{F} = \bigvee_{m=1}^{\infty} \mathcal{F}_m$. We have

$$h_\mu(f) \geq h_\mu(f, \mathcal{P}) = H_\mu\left(\mathcal{P} \middle| \bigvee_{n=1}^{\infty} f^{-n}(\mathcal{P})\right) = H_\mu\left(\mathcal{P} \middle| f^{-1}\left(\bigvee_{n=0}^{\infty} f^{-n}(\mathcal{P})\right)\right)$$

and also, by Proposition 3.1(iii),

$$H_\mu\left(\mathcal{P} \middle| f^{-1}\left(\bigvee_{n=0}^{\infty} f^{-n}\mathcal{P}\right)\right) = \lim_{m\to\infty} H_\mu\left(\mathcal{P} \middle| f^{-1}\left(\bigvee_{n=0}^{m-1} f^{-n}\mathcal{P}\right)\right)$$

and

$$H_\mu(\mathcal{P}|\mathcal{F}) = \lim_{m\to\infty} H_\mu(\mathcal{P}|\mathcal{F}_m).$$

Further,

$$H_\mu\left(\mathcal{P} \middle| f^{-1}\left(\bigvee_{n=0}^{m-1} f^{-n}(\mathcal{P})\right)\right) \geq H_\mu(\mathcal{P}|\mathcal{F}_m).$$

by Proposition 3.1(ii) and (15). Thus

$$h_\mu(f) \geq H_\mu(\mathcal{P}|\mathcal{F}). \tag{17}$$

Let

$$w_m = \sum_{i=1}^{k} 1_{A_i} E_\mu(1_{A_i}|\mathcal{F}_m)$$

for every m and

$$w = \sum_{i=1}^{k} 1_{A_i} E_\mu(1_{A_i}|\mathcal{F}).$$

Here we have rewritten $\mathcal{P} = \{A_1, A_2, \dots, A_k\}$, thus $-\log w_m = I(\mathcal{P}|\mathcal{F}_m)$, $-\log w = I(\mathcal{P}|\mathcal{F})$. Recall that $0 \leq w \leq 1$ μ-a.e. and

$$H_\mu(\mathcal{P}|\mathcal{F}) = \int_I I(\mathcal{P}|\mathcal{F}) \, d\mu \geq \int_K -\log w \, d\mu.$$

Further, take into account that $(I(\mathcal{P}|\mathcal{F}_m))$ converges μ-a.e. to $I(\mathcal{P}|\mathcal{F})$ (hence (w_m) converges μ-a.e. to w) as $m \to \infty$ by Proposition 3.1(iii). Then, in view of (17), to get (12) and (13) it suffices to prove

$$\lim_{m\to\infty} w_m(x) = \frac{\rho(x)}{|f'(x)|\rho(f(x))} \quad \text{for } \mu\text{-a.e. } x \in K.$$

If $x \in I$, then we denote by $A(x)$ and $P_m(x)$, respectively, the elements from the partitions $\mathcal{P}$ and $\mathcal{P}_m$ containing x. Since

$$w_m(x) = E_\mu(1_{A(x)}|\mathcal{F}_m)(x) = \frac{\mu(A(x) \cap f^{-1}(P_m(f(x))))}{\mu(f^{-1}(P_m(f(x))))},$$

we are left to prove that

$$\lim_{m\to\infty} \frac{\mu(A(x) \cap f^{-1}(P_m(f(x))))}{\mu(f^{-1}(P_m(f(x))))} = \frac{\rho(x)}{|f'(x)|\rho(f(x))} \quad \text{for } \mu\text{-a.e. } x \in K. \tag{18}$$

Property (16) implies that $\bigvee_{m=1}^{\infty} \mathcal{P}_m = \mathcal{B}$. Therefore we can apply [27, Theorem 5.1, p. 7] to conclude that if the λ-integrable map p is given, then

$$\frac{\int_{P_m(y)} p\, d\lambda}{\lambda(P_m(y))} \to p(y) \text{ as } m \to \infty \quad \text{for } \lambda\text{-a.e. } y \in I$$

and hence

$$\frac{\int_{P_m(f(x))} p\, d\lambda}{\lambda(P_m(f(x)))} \to p(f(x)) \text{ as } m \to \infty \quad \text{for } \lambda\text{-a.e. } x \in K \tag{19}$$

(use also Lemma 3.11(i)). If $A = K_j$ for some j, then we can extend $f|_A$ to a C^1-diffeomorphism (recall Theorem 3.13(ii)), reason as in the proof of Lemma 3.12, and get

$$\mu(A \cap f^{-1}(P_m(f(x)))) = \int_{P_m(f(x))} 1_{f(A)} \cdot \left(\frac{\rho}{|f'|} \circ (f|_A)^{-1} \right) d\lambda$$

for every m and x. Thus we can apply (19) and assume, for λ-a.e. $x \in K$, that if $\epsilon_{j,m}(x)$ is such that

$$\mu(A \cap f^{-1}(P_m(f(x)))) = \lambda(P_m(f(x))) \left(\epsilon_{j,m}(x) + 1_{f(A)}(f(x)) \cdot \left(\frac{\rho}{|f'|} \circ (f|_A)^{-1} \right)(f(x)) \right), \tag{20}$$

then $(\epsilon_{j,m}(x))$ tends to zero as $m \to \infty$.

We are ready to prove (18). If $x \in K$, then $f(x) \in I \setminus f(N)$, which implies $P_m(f(x)) \subset I \setminus f(N)$ by (14). Thus $f^{-1}(P_m(f(x))) \subset I \setminus N = \bigcup_j K_j$ and we can use (20) to get, for λ-a.e. $x \in K$ (hence for μ-a.e. $x \in K$) and the corresponding $j(x)$ satisfying $A(x) = K_{j(x)}$,

$$\mu(A(x) \cap f^{-1}(P_m(f(x)))) = \lambda(P_m(f(x))) \left(\epsilon_{j(x),m}(x) + \frac{\rho(x)}{|f'(x)|} \right)$$

and

$$\begin{aligned}\mu(f^{-1}(P_m(f(x)))) &= \sum_{j=1}^{r} \mu(K_j \cap f^{-1}(P_m(f(x)))) \\ &= \lambda(P_m(f(x))) \\ &\cdot \left(\sum_{\{j: f(x)\in f(K_j)\}} \epsilon_{j,m}(x) + 1_{f(K_j)}(f(x)) \cdot \left(\frac{\rho}{|f'|} \circ (f|_{K_j})^{-1} \right)(f(x)) \right) \\ &= \lambda(P_m(f(x))) \left(\epsilon_m(x) + \sum_{\{y: f(y)=f(x)\}} \frac{\rho(y)}{|f'(y)|} \right) \\ &= \lambda(P_m(f(x)))(\epsilon_m(x) + \rho(f(x))),\end{aligned}$$

with $(\epsilon_m(x))_m$ tending to zero; in the last equality Theorem 3.13(iii) has been used as well. From this, (18) follows. □

Lemma 7.4. *We have*

$$\int_I (\log|f'| - \log\rho + \log\rho\circ f)\, d\mu = \int_I \log|f'|\, d\mu = \int_I \Lambda_f\, d\mu.$$

Moreover, the Lyapunov exponent $\Lambda_f(x)$ is nonnegative for μ-a.e. $x \in I$.

Proof. Birkhoff's ergodic theorem implies the second inequality and also that, to prove the first one, it suffices to show

$$g^*(x) = \lim_{n\to\infty} \frac{\log|(f^n)'(x)|}{n} - \frac{\log(\rho(x))}{n} + \frac{\log(\rho(f^n(x)))}{n} = \lim_{n\to\infty} \frac{\log|(f^n)'(x)|}{n} = \Lambda_f(x) \quad \text{for } \mu\text{-a.e. } x \in I. \tag{21}$$

Moreover, due to Lemma 7.2, we have that $0 \le g^*(x) < \infty$ for μ-a.e. x. Thus, if (21) holds, then Lemma 7.4 follows.

We must prove (21). The set of points at which ρ vanishes has zero μ-measure, hence $(\log\rho/n)$ tends to zero μ-a.e. Since the set of points x such that $g^*(x) = \infty$ has zero measure, the sequence $(\log(\rho\circ f^n)/n)$ converges μ-a.e. Then it suffices to show that some subsequence of $(\log(\rho\circ f^n)/n)$ goes to zero.

Since $(\log\rho/n)$ tends to zero μ-a.e., it also tends to zero in measure. This means that, for every fixed $\epsilon > 0$, we have

$$\lim_{n\to\infty} \mu(\{|\log\rho|/n > \epsilon\}) = 0,$$

(where $\{|\log\rho|/n > \epsilon\}$ denotes the set of points y satisfying $|\log\rho(y)|/n > \epsilon$). Since μ is invariant and

$$\{|\log(\rho\circ f^n)| > n\epsilon\} = f^{-n}(\{|\log\rho| > n\epsilon\}),$$

we have

$$\mu(\{|\log(\rho\circ f^n)| > n\epsilon\}) = \mu(\{|\log\rho| > n\epsilon\}).$$

Then

$$\lim_{n\to\infty} \mu_x(\{|\log(\rho\circ f^n)|/n > \epsilon\}) = 0.$$

Hence $(\log(\rho\circ f^n)/n)$ also converges to zero in measure, which is well known to imply the existence of a subsequence $(\log|(f^{n_s})'|/n_s)$ converging μ-a.e. to zero as we desired to prove. □

Proof of Theorem D. It follows immediately from Theorem 7.1 and Lemmas 7.2 and 7.4. □

8. **Proof of Theorem E.** This statement is a relatively straightforward consequence of a number of deep but well-known facts. Below we say that a compact interval $K \subset I$ is *totally transitive* if there is a positive integer r (the *period* of K) such that $f^r(K) = K$, the intervals $f^i(K)$, $0 \leq i < r$, have pairwise disjoint interiors, and every map f^{mr}, $m \geq 1$, has a dense orbit in K.

Let P be the set of periodic points of f. As it is proved in [25, Theorem 2(i)], there are finitely many Cantor sets $\{C_j\}$ and totally transitive intervals $\{J_k\}$ with the property that the set of limit points of $(f^n(x))$ is contained in $A = P \cup \bigcup_j C_j \cup \bigcup_k J_k$ for λ-a.e. $x \in I$ and hence, because μ is absolutely continuous, for μ-a.e. $x \in I$. Since μ is invariant, the set of recurrent points has full μ-measure. Then $\mu(A) = 1$.

Observe that if P_r is the set of fixed points of f^r and $p \in P_r$, then p cannot be sensitive unless it is a one-sided or two-sided isolated point of P_r. In particular, the subset of sensitive points of P is countable so $\mu(S(f) \cap P) = 0$. Next, in [34, Theorem E(1)] it is demonstrated that the Cantor sets C_j have zero Lebesgue measure, from which $\mu(C_j) = 0$ for every j. Since

$$0 < \mu(S(f)) = \mu(S(f) \cap A) = \mu(S(f) \cap \bigcup_k J_k) \leq \mu(\bigcup_k J_k),$$

we get $\mu(J) > 0$ for some interval $J = J_{k_0}$. Finally [25, Theorem 2(iv)] guarantees that if r is the period of the interval J and $g = f^r|_J$, then g is ergodic (which means that if a Borel set $B \subset J$ satisfies $g^{-1}(B) = B$, then either $\lambda(B) = 0$ or $\lambda(B) = \lambda(J)$).

Suppose $h_\mu(f) = 0$. Then $\mu(g^n(B)) = \mu(B)$ for every $n \in \mathbb{Z}$ and every Borel set $B \subset J$ (Proposition 3.3(iii)). Therefore μ and λ are equivalent in J. Indeed, if $\mu(B) = 0$ but $\lambda(B) > 0$, and we consider $C = \bigcup_{n,m\geq 0} g^{-m}(g^n(B))$, then $g^{-1}(C) = C$. Hence $\lambda(C) = \lambda(J)$, that is, $\lambda(J \setminus C) = 0$, by the ergodicity of g. Since on the other hand $\mu(C) = 0$ and $\mu(J) > 0$, this contradicts the absolute continuity of μ.

Since J is totally transitive g cannot be monotone, so there are pairwise disjoint intervals U and V such that $g(U) = g(V) = W$. Now

$$\mu(U) = \mu(g(U)) = \mu(W) = \mu(g^{-1}(W)) \geq \mu(U) + \mu(V) > \mu(U)$$

(the last inequality follows from the equivalence of μ and λ in J). We have arrived at a contradiction.

9. **Appendix.** This last section of the paper includes precise formulations and proofs of a number of facts stated in the introductory sections.

Proposition 9.1. *There is a continuous map $f : I \to I$ such that $S(f) = I$ and it does not have full sensitivity to initial conditions.*

Proof. Just take a map having full sensitivity for which the endpoints of the interval are fixed points (say, for instance, the piecewise affine map consisting of three pieces of constant slopes 3, -3 and 3) and copy its graph in each of the squares $[(n+4)/(2n+4), (n+3)/(2n+2)] \times [(n+4)/(2n+4), (n+3)/(2n+2)]$, $n \geq 1$, and $[0, 1/2] \times [0, 1/2]$. See Figure 5. □

Proposition 9.2. *There is a piecewise affine map $f : I \to I$ having full, but no strong, sensitivity to initial conditions.*

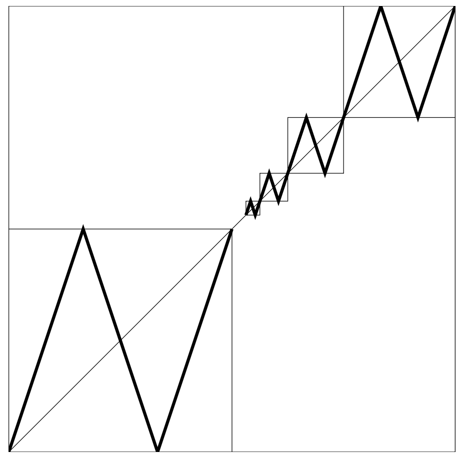

FIGURE 5. For this map all points are sensitive but it does not have full sensitivity to initial conditions.

Proof. Fix an irrational number $0 < \theta < 1$ and consider the map f defined by $f(x) = \text{frac}(x + \theta)$, where $\text{frac}(y)$ denotes the fractional part of y. □

Proposition 9.3. *If $f : I \to I$ is piecewise monotone and $S(f) = I$, then f has full sensitivity to initial conditions.*

Proof. We claim that f has no homtervals. The reason is the following. For such an interval J, we have that either $\lambda(f^n(J)) \to 0$ as $n \to \infty$ (which leads to a contradiction for no interior point of J is sensitive), or there are some r and s such that $K = \bigcup_{m=0}^{\infty} f^{rm+s}(J)$ is an interval which is mapped continuously and non-decreasingly into itself by f^{2r}. Since K can then contain at most a countable number of sensitive points (for they must be one-sided or two-sided isolated fixed points of f^{2r}), we get again a contradiction.

Since f has no homtervals, if the points $0 = a_0 < a_1 < \cdots < a_k = 1$ are such that $f|_{(a_{i-1},a_i)}$ is continuous and strictly monotone for every i and the open interval J is given, then there are some n and j such that $f^n|_J$ is one-to-one continuous and $a_j \in f^n(J)$. Say $a_i \in S_{\delta_i}(f)$, $0 \leq i \leq k$. Then $S_\delta(f) = I$, with $\delta = \min\{\delta_0, \delta_1, \ldots, \delta_k\}/2$. □

Proposition 9.4. *If $f : I \to I$ is continuous and has full sensitivity to initial conditions, then it also has strong sensitivity to initial conditions.*

Proof. The statement is an immediate consequence of the following well-known facts: (a) if $J \subset S_\delta(f)$ for some interval J and some $\delta > 0$, then the closure of $\bigcup_{n=0}^{\infty} f^n(J)$ contains some totally transitive interval (reason as in the proof of [33, Proposition 2.2.5, p. 22] and use [33, Proposition 2.1.17, p. 15]); (b) since two totally transitive intervals have pairwise disjoint interiors (due to [33, Proposition 2.1.10, p. 12]), if f has full sensitivity, then it cannot have infinitely many totally transitive intervals; (c) for every totally transitive interval K there is a number δ such that if J is a subinterval of K, then $\lambda(f^n(J)) > \delta$ for every number n large enough [33, Proposition 2.1.10, p. 12]. □

Proposition 9.5. *There is a quadratic map $f : I \to I$ such that λ-a.e point is asymptotically periodic and $h_\mu(f) > 0$ for some invariant measure μ.*

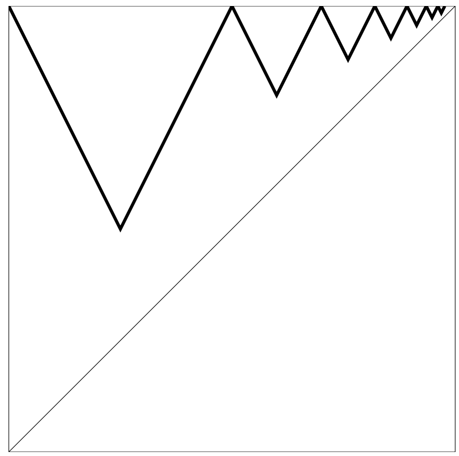

FIGURE 6. A map f such that all but countably many points of I have Lyapunov exponent 2 and all orbits converge to 1.

Proof. Let $f(x) = \beta x(1-x)$, $\beta = 3.83187\ldots$. For this map $1/2$ is a periodic point of period three and hence, according to [16], its orbit attracts that of λ-a.e point of I. On the other hand, f has positive topological entropy ([29]) and hence, by the variational principle [14], admits an invariant measure μ for which $h_\mu(f) > 0$. □

Proposition 9.6. *There is a continuous map $f : I \to I$ such that all orbits converge to the same fixed point and $\lambda(L^+(f)) = 1$.*

Proof. Let f consist of countably many affine pieces of constant slope ± 2 plus the fixed point 1 so that $f(x) > x$ for every $x \in [0, 1)$. For this function all but countably many points of I have Lyapunov exponent 2; nonetheless, all orbits converge to the fixed point 1. See Figure 6. □

Proposition 9.7. *There is a continuous map $f : I \to I$ such that all orbits converge to fixed points and $\lambda(S(f)) = 1$.*

Proof. This map was constructed in [5]. □

Proposition 9.8. *There is a polynomial map $f : I \to I$ with a non-atomic invariant measure μ such that $\mu(S^\mu_\delta(f)) = 1$ for some $\delta > 0$, it has strong sensitivity to initial conditions (with respect to λ) and $h_\mu(f) = 0$.*

Proof. In [9, Theorem 3] a polynomial map $f : I \to I$ having strong sensitivity to initial conditions is constructed with an invariant Cantor ser K that is a wild attractor for f and supports a weakly mixing invariant measure μ. The first implies $h_\mu(f) = 0$ [25, Theorem 4(ii)]; the second means that there is a set N of density zero in $\mathbb{N}$ with the property that if A, B are Borel sets, then $\lim_{N \not\ni n \to \infty} \mu(f^{-n}(A) \cap B) = \mu(A)\mu(B)$ [35, Theorem 1.22, p. 405] (so, in particular, μ is non-atomic). Take $\delta < 2\lambda(\operatorname{conv} K)$. Let $x \in K$ and let B an arbitrary neighbourhood of x. Fix an accumulation point $y \in K$ of the sequence $(f^n(x))_{n \in \mathbb{N} \setminus N}$ and find a small neighbourhood A of one of the endpoints of $\operatorname{conv} K$ so that $\operatorname{dist}(y, A) > \delta$. Since $\mu(f^{-n}(A) \cap B) > 0$ if $n \in N$ is large enough, we conclude that $x \in S^\mu_\delta(f)$. Hence $S^\mu_\delta(f) \supset K$. □

Proposition 9.9. *There is a quadratic map $f : I \to I$ having strong sensitivity to initial conditions and such that $\lambda(L^+(f)) = 0$.*

Proof. According to [21, Theorem 3], if a quadratic map does not admit an acip, then the set of points with positive Lyapunov exponent has zero Lebesgue measure. An example of such a map (having strong sensitivity to initial conditions) was first given in [18]. □

Proposition 9.10. *There is a piecewise affine map $f : I \to I$ for which λ is an invariant measure and such that $h_\lambda(f) = 0$, $L^+(f) = \emptyset$.*

Proof. Use the map from Proposition 9.2. □

REFERENCES

[1] C. Abraham, G. Biau and B. Cadre, *Chaotic properties of mappings on a probability space*, J. Math. Anal. Appl., **266** (2002), 420–431.

[2] C. Abraham, G. Biau and B. Cadre, *On Lyapunov exponent and sensitivity*, J. Math. Anal. Appl., **290** (2004), 395–404.

[3] R. B. Ash, "Real Analysis and Probability," Probability and Mathematical Statistics, No. 11, Academic Press, New York-London, 1972.

[4] Y. Baba, I. Kubo and Y. Takahashi, *Li-Yorke's scrambled sets have measure* 0, Nonlinear Anal., **26** (1996), 1611–1612.

[5] A. Barrio Blaya and V. Jiménez López, *Is trivial dynamics that trivial?*, Amer. Math. Monthly, **113** (2006), 109–133.

[6] A. M. Blokh, *Sensitive mappings of an interval*, Uspekhi Mat. Nauk, **37** (1982), 189–190.

[7] A. Blokh and M. Misiurewicz, *Wild attractors of polymodal negative Schwarzian maps*, Comm. Math. Phys., **199** (1998), 397–416.

[8] A. Boyarsky and P. Góra, "Laws of Chaos. Invariant Measures and Dynamical Systems in One Dimension," Probability and its Applications, Birkhäuser Boston, Inc., Boston, MA, 1997.

[9] H. Bruin, G. Keller and M. St. Pierre, *Adding machines and wild attractors*, Ergodic Theory Dynam. Systems, **17** (1997), 1267–1287.

[10] J. Buzzi, *Thermodynamical formalism for piecewise invertible maps: Absolutely continuous invariant measures as equilibrium states*, in "Smooth Ergodic Theory and its Applications" (Seattle, WA, 1999), Proc. Sympos. Pure Math., **69**, Amer. Math. Soc., Providence, RI, (2001), 749–783.

[11] B. Cadre and P. Jacob, *On pairwise sensitivity*, J. Math. Anal. Appl., **309** (2005), 375–382.

[12] B. D. Craven, "Lebesgue Measure & Integral," Pitman, Boston, MA, 1982.

[13] R. L. Devaney, "An Introduction to Chaotic Dynamical Systems," The Benjamin/Cummings Publishing Co., Inc., Menlo Park, CA, 1986.

[14] E. I. Dinaburg, *A correlation between topological entropy and metric entropy*, (Russian) Dokl. Akad. Nauk SSSR, **190** (1970), 19–22.

[15] E. Glasner and B. Weiss, *Sensitive dependence on initial conditions*, Nonlinearity, **6** (1993), 1067–1075.

[16] J. Guckenheimer, *Sensitive dependence to initial conditions for one-dimensional maps*, Comm. Math. Phys., **70** (1979), 133–160.

[17] F. Hofbauer, *An inequality for the Ljapunov exponent of an ergodic invariant measure for a piecewise monotonic map of the interval*, in "Lyapunov Exponents" (Oberwolfach, 1990), Lecture Notes in Math., **1486**, Springer, Berlin, (1991), 227–231.

[18] S. D. Johnson, *Singular measures without restrictive intervals*, Comm. Math. Phys., **110** (1987), 185–190.

[19] A. Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Inst. Hautes Études Sci. Publ. Math., **51** (1980), 137–173.

[20] A. Katok and B. Hasselblatt, "Introduction to the Modern Theory of Dynamical Systems," With a supplementary chapter by Katok and Leonardo Mendoza, Encyclopedia of Mathematics and its Applications, **54**, Cambridge University Press, Cambridge, 1995.

[21] G. Keller, *Exponents, attractors and Hopf decompositions for interval maps*, Ergodic Theory Dynam. Systems, **10** (1990), 717–744.

[22] A. N. Kolmogorov, *A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces* (Russian), Dokl. Akad. Nauk SSSR (N.S.), **119** (1958), 861–864.

[23] F. Ledrappier, *Some properties of absolutely continuous invariant measures on an interval*, Ergodic Theory Dynam. Systems, **1** (1981), 77–93.

[24] E. N. Lorenz, *The predictability of hydrodynamic flow*, Trans. New York Acad. Sci., Ser. 2, **25** (1963), 409–432.

[25] M. Lyubich, *Ergodic theory for smooth one-dimensional dynamical systems*, Stony Brook preprint, 1991/11, arXiv:math/9201286.

[26] R. Mañé, *Hyperbolicity, sinks and measure in one-dimensional dynamics*, Comm. Math. Phys., **100** (1985), 495–524, Erratum in Comm. Math. Phys., **112** (1987), 721–724.

[27] R. Mañé, "Ergodic Theory and Differentiable Dynamics," Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], **8**, Springer-Verlag, Berlin, 1987.

[28] W. de Melo and S. van Strien, "One-Dimensional Dynamics," Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], **25**, Springer-Verlag, Berlin, 1993.

[29] M. Misiurewicz, *Horseshoes for mappings of the interval*, Bull. Acad. Polon. Sci. Sér. Sci. Math., **27** (1979), 167–169.

[30] W. Parry, "Entropy and Generators in Ergodic Theory," W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[31] V. A. Rohlin, *Exact endomorphisms of a Lebesgue space*, (Russian) Izv. Akad. Nauk SSSR Ser. Mat., **25** (1961), 499–530.

[32] D. Ruelle, *An inequality for the entropy of differentiable maps*, Bol. Soc. Brasil. Mat., **9** (1978), 83–87.

[33] S. Ruette, *Chaos for continuous interval maps. A survey of relationship between the various sorts of chaos*, preprint, Université Paris-Sud, 2003. Available from: http://www.math.u-psud.fr/~ruette/publications.html.

[34] S. van Strien and E. Vargas, *Real bounds, ergodicity and negative Schwarzian for multimodal maps*, J. Amer. Math. Soc., **17** (2004), 749–782. Erratum in J. Amer. Math. Soc., **20** (2007), 267–268.

[35] P. Walters, "An Introduction to Ergodic Theory," Graduate Texts in Mathematics, **79**, Springer-Verlag, New York-Berlin, 1982.

[36] H. Whitney, *On totally differentiable and smooth functions*, Pacific J. Math., **1** (1951), 143–159.

Received September 2010; revised July 2011.

E-mail address: alejobar@um.es
E-mail address: vjimenez@um.es