Schwarzian Derivatives and Cylinder Maps

Araceli Bonifant and John Milnor

Abstract

We describe the way in which the sign of the Schwarzian derivative for a family of diffeomorphisms of the interval I affects the dynamics of an associated skew product map of the cylinder $(\mathbb{R}/\mathbb{Z}) \times I$.

Keywords: asymptotic distribution, attractors, intermingled basins, Schwarzian derivative, skew product.

Mathematics Subject Classification (2000): 37F10, 32H50, 32H02.

1 Introduction.

Ittai Kan has described a simple example of a skew product map from the cylinder $(\mathbb{R}/\mathbb{Z}) \times I$ to itself such that the two boundary circles are measure theoretic attractors whose attracting basins are *intermingled*, in the sense that the intersection of any nonempty open set with either basin has strictly positive measure. (See [Kan, 1994].) This note will consist of three variations on the maps which he introduced.

Sections 3 and 4 will describe Kan's example in slightly more generality, emphasizing the importance of *negative Schwarzian derivative*. Section 5 will show that if we substitute *positive Schwarzian derivative* then the behavior will change drastically, and almost all orbits will have a common asymptotic distribution. In the case of *zero Schwarzian derivative*, §6 will prove in some cases (and conjecture in others) that almost all orbits spend most of the time extremely close to one of the two cylinder boundaries; but that each such orbit passes from the ϵ-neighborhood of one boundary circle to the ϵ-neighborhood of the other infinitely many times on such an irregular schedule that there is no asymptotic measure.

Many technical details are relegated to the two appendices.

2 Preliminaries.

Let $I = [0, 1]$, and let $\mathcal{C}$ be the cylinder $(\mathbb{R}/\mathbb{Z}) \times I$ with boundaries $\mathcal{A}_0 = (\mathbb{R}/\mathbb{Z}) \times \{0\}$ and $\mathcal{A}_1 = (\mathbb{R}/\mathbb{Z}) \times \{1\}$. Let $F : \mathcal{C} \to \mathcal{C}$ be a C^3-differentiable map of the form

$$F(x, y) = \big(kx,\ f_x(y)\big)\,, \tag{1}$$

where $k \geq 2$ is a fixed integer, and where each $f_x : I \to I$ is a diffeomorphism with $f_x(0) = 0$ and $f_x(1) = 1$. Thus the derivative

$$f'_x(y) = \partial f_x(y)/\partial y$$

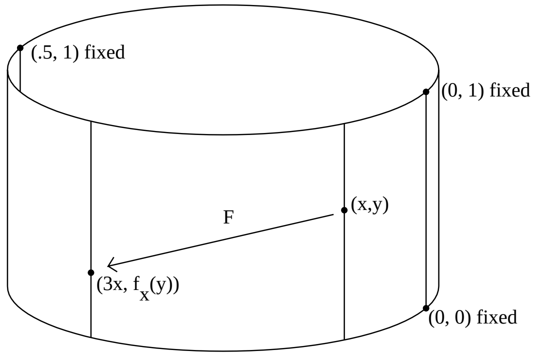

Fig. 1: *The cylinder map F in the case* $k = 3$.

must be strictly positive everywhere.

We next introduce two key concepts that will be needed.

LEMMA A.1. *For ι equal to zero or one, let $\mathcal{B}_\iota$ be the attracting basin of the circle $\mathcal{A}_\iota$. If the transverse Lyapunov exponent*

$$\mathrm{Lyap}(\mathcal{A}_\iota) = \int_{\mathbb{R}/\mathbb{Z}} \log\big(f'_x(\iota)\big)\,dx \tag{2}$$

is negative, then the basin $\mathcal{B}_\iota$ has strictly positive measure. In fact, for almost every $x \in \mathbb{R}/\mathbb{Z}$ the basin $\mathcal{B}_\iota$ intersects the "fiber" $x \times I$ in an interval of positive length. On the other hand, if $\mathrm{Lyap}(\mathcal{A}_\iota) > 0$ then $\mathcal{B}_\iota$ has measure zero.

The proof will be given in Appendix A.

In fact, whenever $\mathrm{Lyap}(\mathcal{A}_\iota) < 0$ it is not hard to see that the circle $\mathcal{A}_\iota$ is a *measure attractor*. By this we mean that it satisfies the following two conditions:

1. $\mathcal{A}_\iota$ is a *minimal measure attracting set*, that is, it has an attracting basin of positive measure, but no closed proper subset has a basin of positive measure.

2. Furthermore, $\mathcal{A}_\iota$ contains a dense orbit, and hence cannot be expressed as the union of strictly smaller closed invariant sets.[1]

Recall that the *Schwarzian derivative* of an interval C^3-diffeomorphism f is defined by the formula

$$\mathcal{S}f(y) \;=\; \frac{f'''(y)}{f'(y)} - \frac{3}{2}\left(\frac{f''(y)}{f'(y)}\right)^2 . \tag{3}$$

We will make a particular study of maps $F(x, y) = \big(kx,\, f_x(y)\big)$ such that the Schwarzian $\mathcal{S}f_x(y)$ has constant sign for almost all $(x, y) \in \mathcal{C}$. Maps f_x with $\mathcal{S}(f_x) < 0$ almost everywhere have

[1]Compare "www.scholarpedia.org/article/Attractor". The following example shows that Condition **2** does not follow from Condition **1**. Consider a flow in the plane such that all orbits near infinity spiral in towards a figure eight-curve, while the open set bounded by either lobe of the figure-eight contains an attracting equilibrium point which attracts all orbits within that open set. Then the figure-eight is a minimal measure attracting set with no dense orbit.

the basic property of *increasing* the cross-ratio $\rho(y_0, y_1, y_2, y_3)$ for all $y_0 < y_1 < y_2 < y_3$ in the interval. (See Appendix B.) Similarly, if $\mathcal{S}(f_x) > 0$ (or if $\mathcal{S}(f_x) \equiv 0$), then f_x will decrease (or will preserve) all such cross-ratios.

The difference between positive, zero and negative Schwarzian may seem somewhat subtle. (Compare Figure 2.) However, we will see that it has a profound influence on the dynamics of the maps.

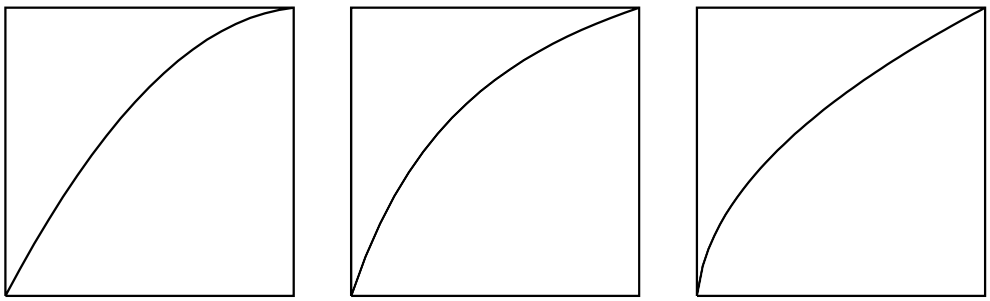

Fig. 2: *Graphs of functions with $\mathcal{S} < 0$, $\mathcal{S} \equiv 0$, and $\mathcal{S} > 0$. The first is the graph of a quadratic polynomial diffeomorphism of the interval.* (*Compare Example 4.4—It follows immediately from Equation* (3) *that $\mathcal{S}f < 0$.*) *The second is the graph of a fractional linear diffeomorphism of the interval, and the third is the graph of the inverse of a quadratic polynomial diffeomorphism.*

3 Negative Schwarzian.

LEMMA 3.1. *If $\mathcal{S}f_x(y)$ has constant sign* (*positive, negative or, zero*) *for almost all (x, y), then $\mathrm{Lyap}(\mathcal{A}_0) + \mathrm{Lyap}(\mathcal{A}_1)$ has the same sign. In particular, if $\mathcal{S}f_x(y) < 0$ for almost all (x, y), then*

$$\mathrm{Lyap}(\mathcal{A}_0) + \mathrm{Lyap}(\mathcal{A}_1) < 0\,, \tag{4}$$

hence at least one of the two boundaries has a basin of positive measure.

Proof. Lemma B.5 (in Appendix B) will show that $f'_x(0)f'_x(1) < 1$ whenever f_x has negative Schwarzian. Integrating the logarithm of this inequality over $\mathbb{R}/\mathbb{Z}$, the inequality (4) follows. Thus the transverse Lyapunov exponent is negative for at least one of the two boundaries. Hence the associated basin has positive measure by Lemma A.1. □

THEOREM 3.2. *If $\mathcal{S}f_x(y) < 0$ almost everywhere, and if both basins have positive measure,*[2] *then there is an almost everywhere defined measurable function $\sigma : \mathbb{R}/\mathbb{Z} \to I$ such that*

$$\begin{array}{lll} (x,y) \in \mathcal{B}_0 & \textit{whenever} & y < \sigma(x)\,, \\ (x,y) \in \mathcal{B}_1 & \textit{whenever} & y > \sigma(x)\,. \end{array}$$

It follows easily that the union $\mathcal{B}_0 \cup \mathcal{B}_1$ has full measure.

[2] We don't know whether this hypothesis is necessary.

In fact we will usually consider maps $f_x(y)$ for which the behavior of F near the two boundaries is similar enough so that $\mathrm{Lyap}(\mathcal{A}_0)$ and $\mathrm{Lyap}(\mathcal{A}_1)$ are equal to each other (or at least have the same sign). For such maps, the condition $\mathcal{S} f_x < 0$ will guarantee that both attracting basins have positive measure.

Proof of Theorem 3.2. Since each f_x is an orientation preserving homeomorphism, there are unique numbers $0 \le \sigma_0(x) \le \sigma_1(x) \le 1$ defined by the property that the orbit of (x, y):

$$\begin{array}{rll} \text{converges to } \mathcal{A}_0 & \text{if} & y < \sigma_0(x) \\ \text{converges to } \mathcal{A}_1 & \text{if} & y > \sigma_1(x) \\ \text{does not converge to either circle} & \text{if} & \sigma_0(x) < y < \sigma_1(x)\,. \end{array}$$

Thus, the area of $\mathcal{B}_0$ can be defined as $\int \sigma_0(x)\,dx$. Since this is assumed to be positive, it follows that the set of all $x \in \mathbb{R}/\mathbb{Z}$ with $\sigma_0(x) > 0$ must have positive measure. In fact, the evident identity $\sigma_0(kx) = f_x\big(\sigma_0(x)\big)$ implies that this set is fully invariant under the ergodic map $x \mapsto kx$. Hence it must actually have full measure. Similarly, the set of x with $\sigma_1(x) < 1$ must have full measure.

We will make use of the property that a map f_x of negative Schwarzian derivative increases the cross-ratio

$$\rho(0,\, y_1,\, y_2,\, 1) \;=\; \frac{y_2\,(1-y_1)}{y_1\,(1-y_2)}\,,$$

that is:

$$\rho\big(0,\, f_x(y_1),\, f_x(y_2),\, 1\big) \;>\; \rho(0,\, y_1,\, y_2,\, 1) \;>\; 1 \qquad \text{for all} \qquad 0 < y_1 < y_2 < 1\,.$$

(See Lemma B.3.) Suppose that the inequalities $0 < \sigma_0(x) < \sigma_1(x) < 1$ were true for a set of $x \in \mathbb{R}/\mathbb{Z}$ of positive Lebesgue measure, then the function

$$r(x) \;=\; \rho\big(0,\, \sigma_0(x),\, \sigma_1(x),\, 1\big) \;\ge\; 1$$

would satisfy $r(kx) > r(x)$ on a set of positive measure, with $r(kx) \ge r(x)$ everywhere. It would follow that

$$\int_{\mathbb{R}/\mathbb{Z}} \frac{dx}{r(kx)} \;<\; \int_{\mathbb{R}/\mathbb{Z}} \frac{dx}{r(x)}\,.$$

But this is impossible: Lebesgue measure is invariant under push-forward by the map $x \mapsto kx$, and it follows that $\int \phi(kx)\,dx = \int \phi(x)\,dx$ for any bounded measurable function ϕ. This contradiction proves that we must have $\sigma_0(x) = \sigma_1(x)$ almost everywhere; and we define $\sigma(x)$ as this common value. □

Remark 3.3. We can then define the *separating measure* β on $\mathcal{C}$ to be the push-forward, under the section, $x \mapsto (x,\, \sigma(x))$, of the Lebesgue measure λ_x, on $\mathbb{R}/\mathbb{Z}$. Evidently β is an F-invariant ergodic probability measure which in some sense describes the "boundary" between the two basins. Since $0 < \sigma(x) < 1$ almost everywhere, it follows easily that both boundaries have measure $\beta(\mathcal{A}_\iota) = 0$.

Remark 3.4. Theorem 3.2 is essentially probabilistic in nature. The same statement would be true in much greater generality: The circle $\mathbb{R}/\mathbb{Z}$ could be replaced by any probability space $\mathbf{X}$, with the k-tupling map replaced by any ergodic measure-preserving transformation $\mathbf{k} : \mathbf{X} \to \mathbf{X}$, and with the correspondence $x \mapsto f_x$ replaced by any measurable transformation from $\mathbf{X}$ to a compact subset of the space of orientation preserving C^3-diffeomorphisms of the interval. The proof would go through without essential change.

As examples, we could equally well replace the expanding map $x \mapsto kx$ by an irrational rotation of the circle, or by a hyperbolic diffeomorphism of the torus. Another example could be obtained by taking the successive $x_i \in \mathbb{R}/\mathbb{Z}$ to be independent random variables, taking our probability space $\mathbf{X}$ to be the cartesian product $(\mathbb{R}/\mathbb{Z})^{\mathbb{N}}$ of countably many copies of the circle with the cartesian product measure. (Compare Hypothesis 6.1 in §6.)

4 Intermingled Basins.

Now assume the following.

Hypothesis 4.1. *There exist angles x^- and x^+ in $\mathbb{R}/\mathbb{Z}$, both fixed under multiplication by k, such that $f_x(y) < y$ for all $0 < y < 1$ and all x in a neighborhood of x^-, and such that $f_x(y) > y$ for all $0 < y < 1$ and all x near x^+.*

It follows that the entire vertical line segment $\{x^-\} \times [0, 1)$ is contained in the basin $\mathcal{B}_0$, and that the entire segment $\{x^+\} \times (0, 1]$ is contained in the basin $\mathcal{B}_1$.

THEOREM 4.2. (Intermingled Basins). *If Hypothesis 4.1 is satisfied, and if both basins have positive measure, then the two basins are intermingled. That is, for every non-empty open set $U \subset \mathcal{C}$, both intersections $\mathcal{B}_0 \cap U$ and $\mathcal{B}_1 \cap U$ have strictly positive measure.*

Proof. Define measures μ_0 and μ_1 on the cylinder by setting $\mu_\iota(S)$ equal to the Lebesgue measure of the intersection $\mathcal{B}_\iota \cap S$ for ι equal to zero or one and for any measurable set S. Clearly the *support* $\mathbf{supp}(\mu_\iota)$, that is the smallest closed set which has full measure under μ_ι, is fully F-invariant. We must prove that this support is equal to the entire cylinder.

To begin, choose any point $(x_0, y_0) \in \mathbf{supp}(\mu_0)$ with $0 < y_0 < 1$. Construct a backward orbit

$$\cdots \mapsto (x_{-2}, y_{-2}) \mapsto (x_{-1}, y_{-1}) \mapsto (x_0, y_0)$$

under F by induction, letting each $x_{-(k+1)}$ be that preimage of x_{-k} which is closest to x^-. Then it is not difficult to see that this backwards sequence converges to the point $(x^-, 1)$. Since $\mathbf{supp}(\mu_0)$ is closed and F-invariant, it follows that $(x^-, 1) \in \mathbf{supp}(\mu_0)$. But the iterated pre-images of $(x^-, 1)$ are everywhere dense in the upper boundary circle $\mathcal{A}_1$, so $\mathcal{A}_1$ is contained in $\mathbf{supp}(\mu_0)$. Since the basin $\mathcal{B}_0$ is a union of vertical line segments $x \times \big[0, \sigma_0(x)\big)$ or $x \times \big[0, \sigma_0(x)\big]$, it follows easily that $\mathbf{supp}(\mu_0)$ is the entire cylinder.

The proof for μ_1 is completely analogous. □

Remark 4.3. In place of a fixed point on the circle, we could equally well use a periodic point $k^p x \equiv x \pmod{\mathbb{Z}}$. It is only necessary to check that the iterated map $F^{\circ p}$ satisfies the required hypothesis.

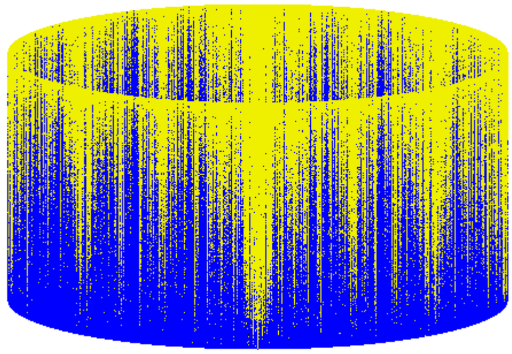

Fig. 3: *Intermingled basins for the cylinder map F of Example 4.4.*

Example 4.4. Following [Kan, 1994], let

$$q_a(y) = y + ay(1-y)\,. \tag{5}$$

If $|a| < 1$, then q_a maps the unit interval diffeomorphically onto itself, with $q_a(0) = 0$ and $q_a(1) = 1$. It is easy to check that $\mathcal{S}q_a(y) < 0$ whenever $a \neq 0$. It then follows from Lemma B.3 that $q_a(y)$ has the property of increasing cross-ratios.

Choose $0 < \epsilon < 1$, and let $p(x) = \epsilon\cos(2\pi x)$. Then for any $k \geq 3$ the map

$$F(x,y) \;=\; \big(kx,\; f_x(y)\big) \qquad \text{where} \qquad f_x(y) = q_{p(x)}(y)$$

will satisfy Hypothesis 4.1 and also the hypotheses of Theorem 3.2. In fact, we can take $x^+ = 0$, and choose x^- to be a fixed point which lies between 1/3 and 2/3. For example, take

$$x^- = \begin{cases} 1/2\,, & \text{for} \quad k \quad \text{odd,} \\ k/(2k-2)\,, & \text{for} \quad k \geq 4 \quad \text{even.} \end{cases}$$

Thus we obtain explicit examples of maps with intermingled basins. (Compare Fig. 3.)

(In fact this argument will work for $k = 2$ also, using the periodic orbit $1/3 \leftrightarrow 2/3$ in place of a fixed point.)

Remark 4.5. Very similar examples of intermingled basins can be observed in rational maps of the projective plane. (Compare [Bonifant, Dabija and Milnor, 2006, §6].) It would be very interesting to know to what extent the examples in the following two sections also have analogs among such rational maps.

5 Positive Schwarzian

In this section we continue to study the cylinder maps $F(x,y) = (kx,\; f_x(y))$, but now assume that $\mathcal{S}f_x > 0$ almost everywhere.

THEOREM 5.1. *If $\mathcal{S}f_x(y) > 0$ for almost all (x, y), then at least one of the transverse Lyapunov exponents* $\mathrm{Lyap}(\mathcal{A}_0)$ *and* $\mathrm{Lyap}(\mathcal{A}_1)$ *is strictly positive. If both are strictly positive, then F has an asymptotic measure.*[3] *That is, there is a uniquely defined probability measure ν on the cylinder such that, for Lebesgue almost every orbit $(x_0, y_0) \mapsto (x_1, y_1) \mapsto \cdots$ and for every continuous test function $\chi : \mathcal{C} \to \mathbb{R}$, the time averages*

$$\frac{1}{n}\Big(\sum_{i=0}^{n-1} \chi(x_i,\, y_i)\Big)$$

converge to the space average $\int_{\mathcal{C}} \chi(x, y)\, d\nu(x, y)$ as $n \to \infty$. (Briefly, almost every orbit is uniformly distributed *with respect to ν.) Furthermore, both boundaries of $\mathcal{C}$ have asymptotic measure $\nu(\mathcal{A}_0) = \nu(\mathcal{A}_1)$ equal to zero.*

Thus, under these hypotheses, almost all orbits of F have the same asymptotic distribution.

Outline of the Proof. Since the proof of this theorem will be slightly circuitous, we first outline the main steps.

- First the circle $\mathbb{R}/\mathbb{Z}$ of the previous section will be replaced by the solenoid

  $$\Sigma \;=\; \Sigma_k \;=\; \underset{\leftarrow}{\lim}\,(\mathbb{R}/k^n\mathbb{Z})\ ,$$

  and the many-to-one map F of $(\mathbb{R}/\mathbb{Z}) \times I$ will be replaced by the associated invertible map $\widetilde{F}$ from $\Sigma \times I$ to itself.

- Since $\widetilde{F}^{-1}$ has negative Schwarzian on each fiber, Lemma 3.1 implies that at least one of the two boundaries of $\Sigma \times I$ has negative Lyapunov exponent under the map $\widetilde{F}^{-1}$. If both are negative, then Lemma A.1 implies that both boundaries have basins of positive measure under $\widetilde{F}^{-1}$. Theorem 3.2 (together with Remark 3.4) then asserts that the union of the attracting basins of the two boundaries $\Sigma \times \{0\}$ and $\Sigma \times \{1\}$ under $\widetilde{F}^{-1}$ will have full measure. In fact there is an almost everywhere defined section

  $$\tilde{x} \;\mapsto\; (\,\tilde{x},\, \sigma(\tilde{x})\,) \tag{6}$$

  from Σ to $\Sigma \times I$ which "separates" the two attracting basins.

- There is a standard ergodic invariant probability measure μ_Σ on the solenoid. Pushing it up to the graph of σ under the section (6), we obtain an ergodic invariant probability measure $\widetilde{\nu}$ on $\Sigma \times I$.

- Since almost all points are pushed *away* from the graph of σ by the inverse map $\widetilde{F}^{-1}$, it follows that they are pushed *towards* this graph by the map $\widetilde{F}$. In this way, we see that $\widetilde{\nu}$ is an asymptotic measure for $\widetilde{F}$.

- Finally, we denote by ν the push-forward of $\widetilde{\nu}$ under the projection

  $$\Sigma \times I \to (\mathbb{R}/\mathbb{Z}) \times I.$$

  This will be the required asymptotic measure for the original cylinder map F.

[3] Terms such as: natural measure or physical measure are also used in the literature to denote this type of measure.

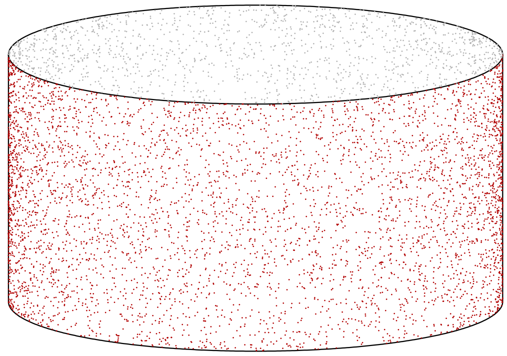

Fig. 4: *50000 points of a randomly chosen orbit for the cylinder map F of Example 5.2.*

Example 5.2. One example of a family of interval diffeomorphisms with positive Schwarzian is given by the inverses

$$f_x(y) \;=\; q_{p(x)}^{-1}(y)\,,$$

where $q_{p(x)}(y)$ is the quadratic map (5) of Example 4.4. Here $\mathcal{S}f_x(y) > 0$ whenever $p(x) \neq 0$. (Compare Proposition B.1 in the appendix.) For this special example, the asymptotic measure ν turns out to be precisely equal to the standard Lebesgue measure λ on the cylinder. In other words:

Randomly chosen orbits are uniformly distributed with respect to Lebesgue measure.

(Compare Fig. 4). To prove this statement, one only needs to show that *Lebesgue measure is F-invariant.* In fact there are k branches of F^{-1} on any small open set $U \subset \mathcal{C}$, each given by

$$F^{-1}(x,y) \;=\; \big(\,x/k,\; y+\epsilon\,\cos(2\pi x/k)y(1-y)\big) \tag{7}$$

for one of the k choices of $x/k \pmod{\mathbb{Z}}$. The Jacobian of this branch (7) is equal to

$$\big(1+\epsilon\,\cos(2\pi x/k)(1-2y)\big)/k\,.$$

Since the sum of $\cos(2\pi x/k) \;=\; \Re e^{2\pi i x/k}$ over the k choices for x/k is zero, the sum of Jacobians is $+1$, which means that F preserves the Lebesgue measure λ. Now Theorem 5.1 asserts that an asymptotic measure exists. Any asymptotic measure ν can necessarily be described as the weak limit of $\big(\lambda + F_*\lambda + \cdots + F_*^{\circ n-1}\lambda\big)/n$ as $n \to \infty$ (using the Lebesgue dominated convergence theorem). In our case, since $F_*\lambda = \lambda$, it follows that ν is precisely equal to λ.

Proof of Theorem 5.1. The argument begins as follows. Denote by Σ the solenoid of backwards orbits

$$\widetilde{x} = \{\cdots \mapsto x_{-2} \mapsto x_{-1} \mapsto x_0\} = \{x_{-n}\}$$

under the map $x \mapsto kx$. Thus Σ maps homeomorphically onto itself under multiplication by k, with the right shift map

$$\{\cdots \mapsto x_{-2} \mapsto x_{-1} \mapsto x_0\} \mapsto \{\cdots \mapsto x_{-3} \mapsto x_{-2} \mapsto x_{-1}\}$$

as inverse. There is a standard probability measure μ_Σ on Σ, defined by the requirement that each projection $\widetilde{x} \mapsto x_{-n}$ is measure preserving.

LEMMA 5.3. $\|\widetilde{x}\|/k \;\le\; \|k\widetilde{x}\| \;\le\; k\,\|\widetilde{x}\|$ *for all* $\widetilde{x}\in\Sigma$.

Proof. This follows easily from the definition. □

Given maps $f_x(y)$ as in §2, we again consider the associated map $F:\mathcal{C}\to\mathcal{C}$. If $\mathcal{S}f_x(y)>0$ almost everywhere, then by the analogue to Lemma 3.1,

$$\mathrm{Lyap}(\mathcal{A}_0)+\mathrm{Lyap}(\mathcal{A}_1)>0.$$

There is also a *natural extension* $\widetilde{F}:\Sigma\times I\;\to\;\Sigma\times I$ of the map F. This is a homeomorphism defined by the formula

$$\widetilde{F}(\widetilde{x},\,y)\;=\;\big(k\,\widetilde{x},\;f_{x_0}(y)\big)\,. \tag{8}$$

LEMMA 5.4. *If* $\mathcal{S}f_x(y)>0$ *for almost all* (x,y)*, and if both* $\mathrm{Lyap}(\mathcal{A}_0)$ *and* $\mathrm{Lyap}(\mathcal{A}_1)$ *are strictly positive, then there exists a measurable function* $\sigma:\Sigma\to I$*, defined almost everywhere, and satisfying the identity*

$$\sigma(k\widetilde{x})=f_{x_0}\big(\sigma(\widetilde{x})\big)$$

for almost all $\widetilde{x}\in\Sigma$*. It follows that the graph of* σ*, that is the set of all pairs*

$$(\widetilde{x},\,\sigma(\widetilde{x}))\;\in\;\Sigma\;\times\;I,$$

is invariant under the extended map $\widetilde{F}:\Sigma\times I\to\Sigma\times I$*, so that* $\widetilde{F}\big(\widetilde{x},\,\sigma(\widetilde{x})\big)=\big(k\widetilde{x},\,\sigma(k\widetilde{x})\big)$ *for almost all* $\widetilde{x}$*.*

Caution: In cases of interest, this function σ will not be continuous and will not be everywhere defined.

Proof of Lemma 5.4. We apply Theorem 3.2 to the inverse map $\widetilde{F}^{-1}$, with the k-tupling map on the circle replaced by the right shift map on the solenoid. This yields a measurable section $\widetilde{x}\;\mapsto\;\big(\widetilde{x},\,\sigma(\widetilde{x})\big)$ from Σ to $\Sigma\times I$. The required F-invariance property then follows easily. □

Next we will show that almost every orbit under $\widetilde{F}$ converges, in a suitable sense, to the graph of σ. Recall that $\sigma(\widetilde{x})$ is well defined and belongs to the open interval $(0,\,1)$ for almost every $\widetilde{x}\in\Sigma$. Thus, for almost every point $(\widetilde{x},\,y)\in\Sigma\times(0,1)$, the quantity

$$r(\sigma(\widetilde{x}),\,y)\;=\;\big|\log\rho\big(0,\,\sigma(\widetilde{x}),\,y,\,1\big)\big|\;\ge\;0$$

is defined and finite, vanishing if and only if $y=\sigma(\widetilde{x})$. We will think of $r(\sigma(\widetilde{x}),\,y)$ as a measure of distance between $\sigma(\widetilde{x})$ and y.

LEMMA 5.5. *Under the same hypothesis, for almost every orbit* $(\widetilde{x}_0,\,y_0)\mapsto(\widetilde{x}_1\,y_1)\mapsto\cdots$ *under* $\widetilde{F}$ *this measure of distance* $r(\sigma(\widetilde{x}_n),\,y_n)$ *converges to zero as* $n\to\infty$*.*

Proof. Since the map $\widetilde{F}(\widetilde{x},\,y)=\big(k\widetilde{x},\,f_{x_0}(y)\big)$ decreases cross-ratios on each fiber (compare Lemma B.3), we have

$$r\big(\sigma(k\widetilde{x}),\,f_{x_0}(y)\big)\;<\;r\big(\sigma(\widetilde{x}),\,y\big)$$

almost everywhere. For any constant $r_0 > 0$, let $N(r_0)$ be the strip consisting of all

$$(\widetilde{x},\, y) \in \Sigma \times (0,1) \qquad \text{with} \qquad r(\sigma(\widetilde{x}),\, y) \;\le\; r_0\,.$$

Evidently $N(r_0)$ is mapped into itself by $\widetilde{F}$. Given constants $0 < r_0 < r_1$, we will also consider the difference set $N(r_1) \smallsetminus N(r_0)$. Let $s(\widetilde{x}) \le 1$ be the supremum of the ratio

$$\frac{r\big(\sigma(k\widetilde{x}),\, f_{x_0}(y)\big)}{r(\sigma(\widetilde{x}),\, y)} \qquad \text{for} \qquad (\widetilde{x},\, y) \in N(r_1) \smallsetminus N(r_0)\,, \tag{9}$$

or in other words for $r_0 < r(\sigma(\widetilde{x}),\, y) \le r_1$. Since the Schwarzian is positive almost everywhere, it is not hard to see that this supremum satisfies $s(\widetilde{x}) < 1$ for almost all $\widetilde{x}$. (Here we make essential use of the fact that $r_0 > 0$, since if the Schwarzian vanishes at $(\widetilde{x},\, \sigma(\widetilde{x}))$ then the ratio (9) would tend to 1 as y tends to $\sigma(\widetilde{x})$.)

Therefore the average of $\log s(\widetilde{x})$ over the solenoid is strictly negative. A straightforward application of the Birkhoff Ergodic Theorem then shows that, for almost every $\widetilde{x}_0 \mapsto \widetilde{x}_1 \mapsto \cdots$, some partial product of the $s(\widetilde{x}_j)$ satisfies

$$s(\widetilde{x}_0) \cdots s(\widetilde{x}_{n-1}) \;<\; r_0/r_1\,.$$

This means that the iterate $\widetilde{F}^{\circ n}$ maps $N(r_1)$ into $N(r_0)$. Since $0 < r_0 < r_1$ can be arbitrary, this completes the proof of Lemma 5.5. □

Now define the probability measure $\widetilde{\nu}$ on $\Sigma \times I$ to be the push-forward of the standard measure μ_Σ on the solenoid under this section $\widehat{\sigma} : \widetilde{x} \mapsto \big(\widetilde{x},\, \sigma(\widetilde{x})\big)$.

LEMMA 5.6. *This $\widetilde{\nu}$ is an asymptotic measure for the extended map $\widetilde{F} : \Sigma \times I \to \Sigma \times I$.*

Proof. We know that almost every orbit $(\widetilde{x}_0,\, y_0) \mapsto (\widetilde{x}_1,\, y_1) \mapsto \cdots$ under $\widetilde{F}$ converges (in the sense of Lemma 5.5) towards the graph of σ. If $\chi : \Sigma \times I \to \mathbb{R}$ is any continuous test function, then it follows easily that the difference between the time averages

$$\Big(\sum_0^{n-1} \chi(\widetilde{x}_i,\, y_i)\Big)/n \qquad \text{and} \qquad \Big(\sum_0^{n-1} \chi\big(\widetilde{x}_i,\, \sigma(\widetilde{x}_i)\big)\Big)/n \;=\; \Big(\sum_0^{n-1} \chi\big(\widehat{\sigma}(\widetilde{x}_i)\big)\Big)/n$$

converges to zero as $n \to \infty$. But the Birkhoff Ergodic Theorem, applied to the bounded measurable function $\chi \circ \widehat{\sigma} : \Sigma \to \mathbb{R}$, asserts that this last time average converges towards the space average

$$\int_\Sigma \chi \circ \widehat{\sigma}(\widetilde{x})\, d\mu_\Sigma(\widetilde{x}) \;=\; \int_{\Sigma \times I} \chi(\widetilde{x},\, y)\, d\widetilde{\nu}(\widetilde{x},\, y)\,,$$

as required. □

Proof of Theorem 5.1 (conclusion). If $\mathcal{S}f_x > 0$ it follows from Lemma B.5 and Lemma A.1 that at least one of the transverse Lyapunov exponents of the boundary circles is strictly positive, hence its corresponding basin has measure zero. If both, $\mathrm{Lyap}(\mathcal{A}_0)$ and $\mathrm{Lyap}(\mathcal{A}_1)$ are strictly positive, then Lemmas 5.4 through 5.6 apply. In this case, both basins $\mathcal{B}_0$ and $\mathcal{B}_1$ have measure zero.

Pushing forward the canonical measure μ_Σ on Σ by the section $\widehat{\sigma} : \widetilde{x} \mapsto (\widetilde{x}, \sigma(\widetilde{x}))$, we obtain an asymptotic measure $\widetilde{\nu} = \widehat{\sigma}_*(\mu_\Sigma)$ for the map $\widetilde{F}$. Now, pushing forward again under the projection

$$(\widetilde{x},\ y)\ \mapsto\ (x_0,\ y)$$

from $\Sigma \times I$ to $(\mathbb{R}/\mathbb{Z}) \times I = \mathcal{C}$, we obtain an F-invariant measure ν on $\mathcal{C}$. Since almost every orbit under $\widetilde{F}$ is uniformly distributed with respect to $\widetilde{\nu}$, it follows that almost every orbit under F is uniformly distributed with respect to ν. □

Remark 5.7. In the spirit of Remark 3.3 one could say, that the *separating measure* for $\widetilde{F}^{-1}$ is an *asymptotic measure* for $\widetilde{F}$.

6 Zero Schwarzian

This section will study the intermediate case where each orientation preserving diffeomorphism $f_x : I \to I$ has Schwarzian $\mathcal{S}f_x$ identically zero. Such a map is necessarily fractional linear, and can be written for example as

$$y\ \mapsto\ \frac{ay}{1+(a-1)y} \qquad \text{with} \qquad a\ >\ 0\,, \tag{10}$$

where a is the derivative at $y = 0$. It will be convenient to replace y by the *Poincaré arclength coordinate*[4]

$$t(y)\ =\ \log\rho(0,\,1/2,\,y,\,1)\ =\ \log\frac{y}{1-y}\,, \tag{11}$$

which varies over the entire real line for $0 < y < 1$, with inverse $y = e^t/(1 + e^t)$.

Since we are assuming $\mathcal{S}f_x$ identically zero, it follows that each f_x preserves cross-ratios or Poincaré distances. (See Equation (11), as well as Remark B4 of Appendix B.) Therefore, in terms of the Poincaré arclength coordinate t, the map f_x will simply be a translation, $t \mapsto t + c$ where c is a constant depending on x. In other words,

$$t(f_x(y)) = t(y)\ +\ c\,, \qquad \text{where} \qquad c\ =\ \log(a)\ \in\ \mathbb{R} \qquad \text{or} \qquad a\ =\ e^c\,.$$

Using this displacement c in place of the original parameter a , the 1-parameter group of fractional linear transformations of the unit interval takes the form

$$g_c(y)\ =\ \frac{e^c y}{1+(e^c-1)y}\,,$$

where $g_{c+c'} = g_c \circ g_{c'}$. Given any bounded measurable function p from $\mathbb{R}/\mathbb{Z}$ to $\mathbb{R}$, we can set $c = p(x)$ to obtain an associated cylinder map

$$F(x,\,y)\ =\ \big(kx,\ g_{p(x)}(y)\big)\,.$$

If we use the coordinate $t \in \mathbb{R}$ in place of $y \in (0,1)$, then this skew product cylinder map F will correspond to a map

$$\widetilde{F}\ :\ (x,\,t)\ \mapsto\ \big(k\,x,\ t + p(x)\big)$$

[4] If we embed the unit interval in the complex open disk of radius 1/2 centered at 1/2, then $|t(y_1) - t(y_0)|$ can be described as the Poincaré distance between y_1 and y_0, or in other words as the distance as measured by the Poincaré metric for this disk. Compare Appendix B.

of $(\mathbb{R}/\mathbb{Z})\times\mathbb{R}$. (The map $\widetilde{F}$ has a canonical invariant measure $dx\,dt$. However, this is not very useful since the total area $\int\int dx\,dt$ is infinite.) The dynamics of F (or of $\widetilde{F}$) under iteration is governed by the average

$$\mathrm{Lyap}(\mathcal{A}_0) \quad = \quad -\,\mathrm{Lyap}(\mathcal{A}_1) \quad = \quad \int_{\mathbb{R}/\mathbb{Z}} p(x)\,dx \tag{12}$$

of the displacement $p(x)$. For almost any orbit $\widetilde{F}:(x_0,\,t_0)\mapsto\cdots\mapsto(x_n,\,t_n)\mapsto\cdots,$ it follows from the Birkhoff Ergodic Theorem that the time average

$$(t_n-t_0)/n \;=\; \big(p(x_0)+\cdots+p(x_{n-1})\big)/n$$

converges to the space average $\mathrm{Lyap}(\mathcal{A}_0)$ as $n\to\infty$. Thus if $\mathrm{Lyap}(\mathcal{A}_0)>0$ then it follows that t_n will converge to $+\infty$. In other words, the corresponding orbit for F will converge towards the upper cylinder boundary $\mathcal{A}_1$, so that $\mathcal{A}_1$ will be a global attractor under F. Similarly, if $\mathrm{Lyap}(\mathcal{A}_0)<0$ then the lower boundary $\mathcal{A}_0$ will be a global attractor. (As in Remark 3.4, this proof would work equally well in the more general case where the k-tupling map on the circle is replaced by any ergodic measure-preserving transformation from a probability space to itself.)

The borderline case where the average (12) is exactly zero, is much more interesting. We have been able to prove precise results only in the very special case of a rather different map, described as follows, in which the successive differences $\Delta t_n=t_{n+1}-t_n$ are bounded identically distributed independent random variables with mean zero. However, we conjecture that the same results would hold in the original cylinder example, as described above.

Hypothesis 6.1. *For the rest of this section, we replace the k-tupling map on the circle by the following ergodic measure preserving transformation on a probability space.* (*Compare Remark 3.4.*) *Starting with some standard probability space X, which we may as well take to be the unit interval with its Lebesgue measure, we form the infinite cartesian product*

$$X^{\mathbb{N}} \;=\; X\times X\times X\times\cdots$$

with the cartesian product measure. Thus a point of $X^{\mathbb{N}}$ is an infinite sequence $\mathbf{x}=(x_0,\,x_1,\,x_2,\,\ldots)$ of points of X, which we think of as identically distributed independent random variables. We will also need a measurable function $x\mapsto f_x$ from X to a suitable compact space of orientation preserving diffeomorphisms of the interval. Here we exclude the trivial case where f_x is the identity map for almost all x. The associated map $\mathbf{F}$ from $X^{\mathbb{N}}\times I$ to itself is then defined by

$$\mathbf{F}\big((x_0,\,x_1,\,x_2,\,\ldots),\;y\big) \quad = \quad \big((x_1,\,x_2,\,x_3,\,\ldots),\;f_{x_0}(y)\big)\,. \tag{13}$$

Thus the k-tupling map on the circle is replaced by the left shift transformation $\sigma: X^{\mathbb{N}}\to X^{\mathbb{N}}$. Note that only the initial entry x_0 of $\mathbf{x}$ affects the image $f_{x_0}(y)$. We will continue to use the notations $\mathcal{A}_1$ and $\mathcal{A}_0$ for the upper and lower boundaries, where now $\mathcal{A}_\iota=X^{\mathbb{N}}\times\iota$ for $\iota\in\{0,1\}$.

Now let us assume that the f_x have Schwarzian identically zero, and let us use the alternate variable $t=t(y)$ as in (11). If $\mathbf{x}$ is randomly chosen, it then follows that the successive steps

$$\Delta t_n \;=\; t_{n+1}-t_n \;=\; p(x_n)$$

are bounded identically distributed independent random variables which do not depend on the value of t_n. In fact Δt_n depends only on the n-th component x_n of $\mathbf{x}$. Switching back from the variable $t \in \mathbb{R}$ to the original variable $y \in [0, 1]$, we want to study the behavior of a typical orbit

$$(\mathbf{x}, y_0) \mapsto (\sigma(\mathbf{x}), y_1) \mapsto (\sigma^{\circ 2}(\mathbf{x}), y_2) \mapsto \cdots . \tag{14}$$

Let $\overline{y}_n = (y_0 + y_1 + \cdots + y_{n-1})/n \in [0,1]$ be the n-th time average for this orbit (14).

THEOREM 6.2. *With* $\mathbf{F}$ *as in* (13)*, if all of the* f_x *have Schwarzian identically zero, and if* $\mathrm{Lyap}(\mathcal{A}_0) = -\mathrm{Lyap}(\mathcal{A}_1) = 0$ *but* $p(x)$ *is not identically zero, then there is no asymptotic measure. In fact, for almost all* $\mathbf{x} \in X^{\mathbb{N}}$ *and for all* $y_0 \in (0, 1)$*, the associated sequence* $\{\overline{y}_n\}$ *of time averages satisfies*

$$\limsup_{n\to\infty} \overline{y}_n = 1 \qquad \textit{and} \qquad \liminf_{n\to\infty} \overline{y}_n = 0 .$$

In other words, for almost all $(\mathbf{x}, y_0)$ and for any $\epsilon > 0$, there are infinitely many values of n such that the orbit, up to time n, has spent "most" of its time with $y_i > 1 - \epsilon$, but there are also infinitely many n for which the orbit has spent most of its time in the neighborhood $y_i < \epsilon$. It follows immediately that there cannot be any asymptotic measure.

Proof of Theorem 6.2. We are indebted to Harry Kesten for substantial help with the argument, which will be based on two classical theorems in probability theory. First we will need the following.

Hewitt-Savage Zero-One Law. *The action of the group of finite permutations of* $\mathbb{N}$ *on the space of sequences* $X^{\mathbb{N}}$ *is ergodic. That is, any measurable subset which is invariant under finite permutation of the coordinates must have measure either zero or one.*

(See for example [Feller **2**, 1966, IV.6].) Now, for any $(\mathbf{x}, y_0) \in X^{\mathbb{N}} \times (0, 1)$, let $L(\mathbf{x}, y_0) \in [0, 1]$ be the lim sup of the associated sequence $\{\overline{y}_n\}$. For any constant L_0, the set

$$\{\mathbf{x} \in X^{\mathbb{N}} \;:\; L(\mathbf{x}, y_0) < L_0\}$$

is clearly invariant under finite permutations of $\mathbb{N}$. Using the Zero-One Law, it follows easily that this lim sup is independent of $\mathbf{x}$ for almost all $\mathbf{x}$. Let us denote this common value by $L(y_0) \in [0, 1]$.

Next we show that this upper limit $L(y_0)$ is independent of $y_0 \in (0, 1)$. It is clear from the construction that $L(y_0) = L\big(f_x(y_0)\big)$ for almost all x. On the other hand, it is also clear that the function $y_0 \mapsto L(y_0)$ is monotone increasing. Since $f_x(y_0)$ will be greater than y_0 with positive probability, and also less than y_0 with positive probability, this proves that the function $y_0 \mapsto L(y_0)$ is locally constant, and hence constant.

To evaluate this constant, we will need the following. Again consider the sequence $\{y_n\}$ where the differences $\Delta y_n(\mathbf{x}) = y_{n+1} - y_n$ are bounded, identically distributed random variable with mean zero but not identically zero.

Arc-Sine Law. *Let* $\psi_n(\mathbf{x})$ *be the step function which takes the value* $+1$ *if* $y_n > y_0$ *and the value zero otherwise, and let* $\overline{\psi}_n(\mathbf{x}) = (\psi_1 + \cdots + \psi_n)/n \in [0,\,1]$ *be its* n*-th time average. Then for any constant* $\alpha \in [0,\,1]$ *the probability that* $\overline{\psi}_n(\mathbf{x}) < \alpha$ *converges to the expression*

$$\frac{2}{\pi}\arcsin\sqrt{\alpha} \qquad \textit{as} \qquad n \to \infty\,. \tag{15}$$

Compare the slightly more general statement in [ERDÖS AND KAC, 1947]. (If we use the variable $t \in \mathbb{R}$ rather than $y \in (0,1)$, then it is easy to check that the inequality $y_n > y_0$ depends only on $\mathbf{x}$ and not on y_0.)

Note that the correspondence $\alpha \mapsto (2/\pi)\arcsin\sqrt{\alpha}$ defines a homeomorphism from the unit interval onto itself. As an immediate consequence, we obtain the following equality.

$$\limsup_{n\to\infty}\ \overline{\psi}_n(\mathbf{x}) \;=\; 1 \qquad \text{for almost all} \quad x\,. \tag{16}$$

In fact, for any $\epsilon > 0$, it follows easily from the Arc-Sine Law that $\limsup \overline{\psi}_n(\mathbf{x}) > 1-\epsilon$ with positive probability. Hence, by the Zero-One Law, this inequality is true with probability one. This proves (16).

On the other hand, the inequality

$$\overline{y}_n \;\geq\; y_0\,\overline{\psi}_n$$

follows easily from the definitions. Taking the lim sup of both sides as $n \to \infty$ and applying (16), we see that $L(y_0) \geq y_0$. Since $L(y_0)$ is independent of y_0, and since y_0 can be arbitrarily close to one, this proves that $L(y_0) = 1$, as required. The proof that $\liminf \overline{y}_n = 0$ for all $y_0 \in (0,\,1)$ and for almost all $\mathbf{x}$ is completely analogous. □

Here is a supplementary statement. It follows from Theorem 6.2 that typical orbits in $X^{\mathbb{N}} \times (0,1)$ spend a great deal of time extremely close to one or the other of the two boundaries $\mathcal{A}_0$ and $\mathcal{A}_1$. They must make the transition from one boundary neighborhood to the other infinitely often. However, the next result says that they spend *most* of the time very close to one boundary or the other.

Let J be a compact interval which is strictly contained in $(0,\,1)$. (For example we could take $J = [\epsilon,\,1-\epsilon]$.) Given an orbit $(\mathbf{x},\,y_0) \mapsto (\sigma(\mathbf{x}),\,y_1) \mapsto \cdots$, let

$$\eta_n^J \;=\; \eta_n^J(\mathbf{x},\,y_0) \;=\; \begin{cases} 1 & \text{if} \qquad y_n \in J, \\ 0 & \text{otherwise}\,. \end{cases}$$

Thus the time average $\overline{\eta}_n^J \;=\; (\eta_0^J + \eta_1^J + \cdots + \eta_{n-1}^J)/n$ measures the fraction of the y_i with $i < n$ which lie in the interval J.

THEOREM 6.3. *Under the hypothesis of Theorem 6.2 and with notations as above, the sequence of time averages* $\overline{\eta}_n^J(\mathbf{x},\,y_0)$ *converges to zero for almost every* $\mathbf{x}$ *and for all* y_0.

We are indebted to M. Lyubich for his help with the following proof, which will depend on a study of random walks on the circle. Consider a random walk

$$X^{\mathbb{N}} \times \mathbb{R}/\mathbb{Z} \;\to\; X^{\mathbb{N}} \times \mathbb{R}/\mathbb{Z} \qquad \text{of the form} \qquad (\mathbf{x},\,t) \;\mapsto\; \big(\sigma(\mathbf{x}),\; t + \check{p}(x_0)\big)\,, \tag{17}$$

where $\breve{p} : X \to \mathbb{R}/\mathbb{Z}$. It will be convenient to simplify this description as follows. Let ξ be the given probability measure on X, let $\mu = \breve{p}_*(\xi)$ be the pushed-forward probability measure on $\mathbb{R}/\mathbb{Z}$, and let $S \subset \mathbb{R}/\mathbb{Z}$ be the support of this measure (the smallest closed subset which has full measure under μ). Then we can describe the random walk as the transformation

$$\breve{F} \;:\; (\mathbb{R}/\mathbb{Z}) \times S^{\mathbb{N}} \;\to\; (\mathbb{R}/\mathbb{Z}) \times S^{\mathbb{N}}$$

which is given by

$$\breve{F}(\tau,\,\mathbf{s}) \;=\; \breve{F}(\tau,\; s_0,\; s_1,\; \ldots) \;=\; (\tau + s_0,\; s_1,\; s_2,\; \ldots)\,. \tag{18}$$

(Here we have interchanged the order of the factors $\mathbb{R}/\mathbb{Z}$ and $S^{\mathbb{N}}$ for convenience, for this proof only.) Evidently this transformation $\breve{F}$ is measure-preserving, using the Lebesgue measure λ on $\mathbb{R}/\mathbb{Z}$ and using the measure $\mu^{\mathbb{N}}$ on $S^{\mathbb{N}}$. If the support S of μ happens to be contained in a finite cyclic subgroup of $\mathbb{R}/\mathbb{Z}$, then every orbit of $\breve{F}$ will project to a finite periodic orbit in $\mathbb{R}/\mathbb{Z}$, hence $\breve{F}$ cannot be ergodic. Conversely, we have the following.

LEMMA 6.4. *If S is not contained in any finite cyclic subgroup of $\mathbb{R}/\mathbb{Z}$, then the formula* (18) *describes an ergodic transformation of $(\mathbb{R}/\mathbb{Z}) \times S^{\mathbb{N}}$.*

Proof. (Compare [Lévy, 1939], [Kawada and Ito, 1940].) We must show that every bounded measurable $\breve{F}$-invariant function $\phi : (\mathbb{R}/\mathbb{Z}) \times S^{\mathbb{N}} \to \mathbb{R}$ must be constant almost everywhere. Given such a function ϕ, let $\phi_0 : \mathbb{R}/\mathbb{Z} \to \mathbb{R}$ be the average over all $\mathbf{s} \in S^{\mathbb{N}}$,

$$\phi_0(\tau) \;=\; \int \phi(\tau,\,\mathbf{s})\, d\mu^{\mathbb{N}}(\mathbf{s})\,.$$

We will first show that ϕ_0 is constant almost everywhere. Let $\widehat{\phi}_0 : \mathbb{Z} \to \mathbb{C}$ be the Fourier transform

$$\widehat{\phi}_0(q) \;=\; \int_{\mathbb{R}/\mathbb{Z}} \phi_0(\tau)\, e^{-2\pi i q \tau}\, d\tau\,.$$

Then the required assertion that ϕ_0 is constant almost everywhere translates to the statement that $\widehat{\phi}_0(q) = 0$ for all integers $q \neq 0$.

Let $\widehat{\mu} : \mathbb{Z} \to \mathbb{C}$ be the Fourier transform

$$\widehat{\mu}(q) \;=\; \int_S e^{-2\pi i q s}\, d\mu(s)\,.$$

Since $\widehat{\mu}(q)$ is a weighted average of points on the unit circle, we must have $|\widehat{\mu}(q)| \leq 1$. Furthermore $\widehat{\mu}(q) \neq 1$ for $q \neq 0$, because of the hypothesis that the support S is not contained in the cyclic group consisting of s with $qs \equiv 0 \pmod{\mathbb{Z}}$. Note the identity

$$\phi_0(\tau) \;=\; \int_S \phi_0(\tau + s)\, d\mu(s)\,, \tag{19}$$

which can be proved by averaging the equation

$$\phi(\tau,\, s_0,\, s_1,\, \ldots) \;=\; \phi\big(\breve{F}(\tau,\, s_0,\, s_1,\, \ldots)\big) \;=\; \phi(\tau + s_0,\; s_1,\; s_2,\; \ldots) \tag{20}$$

first over all choices for $(s_1,\, s_2,\, \ldots)$, and then over s_0.

We take the Fourier transform of this equation (19) by multiplying it by $e^{-2\pi i q\tau}\,d\tau$ and then integrating, using the substitution $t=\tau+s$. This yields

$$\widehat{\phi}_0(q) \;=\; \int e^{-2\pi iqt}\phi_0(t)\,dt \int e^{2\pi iqs}\,d\mu(s) \;=\; \widehat{\phi}_0(q)\,\widehat{\mu}(-q)\,.$$

For $q\neq 0$, since $\widehat{\mu}(-q)\neq 1$, this implies that $\widehat{\phi}_0(q)=0$. Therefore the function $\phi_0(\tau)$ takes some constant value v for almost all τ.

More generally, we can define a function

$$\phi_n : (\mathbb{R}/\mathbb{Z})\times S^n \;\to\; \mathbb{R}$$

by setting $\phi_n(\tau,\, s_0,\, \ldots\, s_{n-1})$ equal to the average of $\phi(\tau,\,\mathbf{s})$ over all choices for $s_n,\, s_{n+1},\, \ldots$. We can compute this function inductively by the formula

$$\phi_n(\tau,\, s_0,\, s_1,\, \ldots,\, s_{n-1}) \;=\; \phi_{n-1}(\tau+s_0,\, s_1,\, s_2\,,\ldots,\, s_{n-1})\,,$$

which can be proved by averaging equation (20) over all choices for $s_n,\, s_{n+1},\, \ldots$. Since we know that ϕ_0 takes the constant value v almost everywhere, *it follows inductively that* ϕ_n *takes this same constant value* v *almost everywhere.*

To prove that $\phi(\tau,\,\mathbf{s})$ also takes value v almost everywhere, we proceed as follows. By an open set $U_n\subset(\mathbb{R}/\mathbb{Z})\times S^{\mathbb{N}}$ *of level* n we will mean the preimage of an open subset $U'_n\subset(\mathbb{R}/\mathbb{Z})\times S^n$ under the natural projection map from infinite product to finite product. Note that the average of ϕ over such a set U_n can be identified with the average of ϕ_n over U'_n, which is clearly equal to the common value v.

On the other hand, the open sets of finite level form a basis for the topology of $(\mathbb{R}/\mathbb{Z})\times S^{\mathbb{N}}$. It follows easily that any open set $U\subset(\mathbb{R}/\mathbb{Z})\times S^{\mathbb{N}}$ can be expressed as the union of an increasing sequence $U_1\subset U_2\subset\cdots$ of open sets of finite level. Hence the average of ϕ over U, being the limit of the averages over the U_n, is also equal to v. Finally, the average of ϕ over any set $\Sigma\subset(\mathbb{R}/\mathbb{Z})\times S^{\mathbb{N}}$ of positive measure is equal to v, since any such Σ has open neighborhoods U with $U\smallsetminus\Sigma$ of arbitrarily small measure. Therefore, the set of $(\tau,\,\mathbf{s})$ with $\phi>v$ or with $\phi<v$ must have measure zero. This completes the proof that the map $\breve{F}$ is ergodic. □

Proof of Theorem 6.3. It will be convenient to work with the coordinate $t=t(y)\in\mathbb{R}$ rather than the original coordinate $y\in(0,\,1)$. Thus J will now denote some (possibly very large) closed interval of real numbers. We will show that for a random walk of the form

$$\widetilde{F}(\mathbf{x},\, t) \;=\; \big(\sigma(\mathbf{x}),\, t+p(x_0)\big)$$

on $X^{\mathbb{N}}\times\mathbb{R}$, the fraction of time $\overline{\eta}^J_n(\mathbf{x},\, y_0)$ spent in the compact interval $J\subset\mathbb{R}$ will tend to zero as $n\to\infty$ for almost all $\mathbf{x}$. Let $L\gg 0$ be a large real number. By projecting the real line to the circle $\mathbb{R}/(L\mathbb{Z})$ (or briefly $\mathbb{R}/L$), we obtain a corresponding random walk $\breve{F}$ on this circle of length L. Note that the interval $J\subset\mathbb{R}$ of length $|J|$ projects to an interval $\breve{J}\subset\mathbb{R}/L$ of the same length. (Here we assume that $L>|J|$.)

For almost every choice of L, it is not hard to see that the push-forward $\breve{p}_*(\xi)$ of the canonical measure on X to a measure μ on $\mathbb{R}/L$ will satisfy the hypothesis of Lemma 6.4. For any orbit $(\mathbf{x},\, t_0)\mapsto(\sigma(\mathbf{x}),\, t_1)\mapsto\cdots$, we can project the sequence of points $t_0,\, t_1,\, \ldots\in\mathbb{R}$ to the circle, thus obtaining a corresponding sequence of points $\tau_0,\, \tau_1,\, \cdots\in\mathbb{R}/L$. Since the map

$$(\mathbf{x},\, \tau) \;\mapsto\; \big(\sigma(\mathbf{x}), \tau+\breve{p}(x_0)\big)$$

is ergodic, it follows from the Birkhoff Ergodic Theorem that the orbit $(\mathbf{x}, \tau_0) \mapsto \big(\sigma(\mathbf{x}), \tau_1\big) \mapsto \cdots$ will be uniformly distributed in $X^{\mathbb{N}} \times (\mathbb{R}/L)$ for almost all $\mathbf{x}$. It follows that the image sequence $\tau_0, \tau_1, \ldots \in \mathbb{R}/L$ will be uniformly distributed around the circle for almost all $\mathbf{x}$ and for all τ_0. In particular, this sequence will visit a given interval $\breve{J} \subset \mathbb{R}/L$ of length $|J|$ with limiting frequency precisely equal to $|J|/L$. Evidently any visit by the lifted sequence $\{t_i\}$ to the interval $J \subset \mathbb{R}$ will give rise to a visit by $\{\tau_i\}$ to $\breve{J}$. Therefore this ratio $|J|/L$ is an upper bound for the frequency of visits of $\{t_i\}$ to J. Since L can be arbitrarily large, it follows that the frequency of visits of $\{t_i\}$ to the interval $J \subset \mathbb{R}$ is zero, as asserted. □

Appendix A: The Transverse Exponent.

Let (x, ι) be any point of the boundary circle $\mathcal{A}_\iota$, where ι can be either 0 or 1. By definition the *transverse Lyapunov exponent* along the circle $\mathcal{A}_\iota$ at (x, ι) is defined by

$$\mathrm{Lyap}_{\mathcal{A}_\iota}(x) = \lim_{k\to\infty} \frac{1}{k} \log\left|\frac{\partial F^{\circ k}}{\partial y}(x,y)\right| \qquad \text{evaluated at} \quad y = \iota\,,$$

whenever this limit exists. (In this case, *transverse* really means *normal*.) Here $F(x,y) = \big(kx,\ f_x(y)\big)$ as usual. By the chain rule, the above expression can be written as

$$\mathrm{Lyap}_{\mathcal{A}_\iota}(x) = \lim_{k\to\infty} \frac{1}{k} \log\Big(f'_{x_0}(\iota) f'_{x_1}(\iota) \cdots f'_{x_{k-1}}(\iota)\Big)$$

where $x_0 \mapsto x_1 \mapsto \cdots$ is the orbit of $x = x_0$.

Let us denote by λ the 2-dimensional Lebesgue measure on the cylinder $\mathcal{C} = (\mathbb{R}/\mathbb{Z}) \times I$, and by λ_x the 1-dimensional Lebesgue measure along $\mathbb{R}/\mathbb{Z}$. Since λ_x is ergodic and invariant under multiplication by k (see Equation (1)), it follows from the Birkhoff Ergodic Theorem that this transverse Lyapunov exponent is defined and independent of x for almost all x, and is equal to the integral

$$\mathrm{Lyap}_{\mathcal{A}_\iota} = \int_{\mathbb{R}/\mathbb{Z}} \log\Big(f'_x(0)\Big) dx,$$

for almost all x.

Let us prove now the Lemma stated in §2.

LEMMA A.1. *For ι equal to zero or one, let $\mathcal{B}_\iota$ be the attracting basin of the circle $\mathcal{A}_\iota$. If the transverse Lyapunov exponent*

$$\mathrm{Lyap}(\mathcal{A}_\iota) = \int_{\mathbb{R}/\mathbb{Z}} \log\big(f'_x(\iota)\big)\, dx \tag{A1}$$

is negative, then the basin $\mathcal{B}_\iota$ has strictly positive measure. In fact, for almost every $x \in \mathbb{R}/\mathbb{Z}$ the basin $\mathcal{B}_\iota$ intersects the "fiber" $x \times I$ in an interval of positive length. On the other hand, if $\mathrm{Lyap}(\mathcal{A}_\iota) > 0$ then $\mathcal{B}_\iota$ has measure zero.

Proof. First consider the case $\iota = 0$ with $\mathrm{Lyap}(\mathcal{A}_0) < 0$. By Taylor's expansion restricted to the fiber over x, we have

$$f_x(y) = f'_x(0)y + O(y^2)\,,$$

for all $(x, y) \in (\mathbb{R}/\mathbb{Z}) \times I$. Choose $K > 0$ so that $f_x(y)$ is uniformly bounded, i.e.,

$$f_x(y) \;\le\; y\big(f'_x(0) + Ky\big) \qquad \text{for} \quad \text{all} \qquad (x, y)\,.$$

For any $\eta > 0$, it follows that

$$f_x(y) \;\le\; y\big(f'_x(0) + \eta\big) \qquad \text{whenever} \qquad y < \frac{\eta}{K}\,. \tag{A2}$$

Since $\mathrm{Lyap}(\mathcal{A}_0) < 0$, we can choose $\eta > 0$ small enough so that

$$\int_{\mathbb{R}/\mathbb{Z}} \log\big(f'_x(0) + \eta\big)dx \;<\; 0\,. \tag{A3}$$

It will be convenient to introduce the notation

$$a(x) \;=\; \log\big(f'_x(0) + \eta\big)\,. \tag{A4}$$

Consider some orbit $(x_0,\, y_0) \mapsto (x_1,\, y_1) \mapsto (x_2,\, y_2) \mapsto \cdots$. By the Birkhoff Ergodic Theorem, the averages

$$\frac{1}{n}\Big(a(x_0) + a(x_1) + \ldots + a(x_{n-1})\Big)$$

converge to $\int_{\mathbb{R}/\mathbb{Z}} a(x)\,dx < 0$ for almost all x_0. In particular, it follows that the n-fold sum

$$A_n(x_0) \;=\; a(x_0) + a(x_1) + \ldots + a(x_{n-1})$$

converges to negative infinity as $n \to \infty$. Hence the maximum

$$A_{\max}(x) \;=\; \max_{n \ge 0} A_n(x)$$

is certainly defined and finite for almost all x, therefore a measurable function. Now suppose that

$$y_0 \;\le\; \frac{\eta}{K}\, e^{-A_{\max}(x_0)}\,. \tag{A5}$$

Then a straightforward induction shows that

$$y_n \;\le\; \frac{\eta}{K} e^{A_n(x_0) - A_{\max}(x_0)} \;\le\; \frac{\eta}{K}$$

for all n. Since $A_n(x_0)$ converges to $-\infty$, it follows that y_n tends to zero, so that $(x_0,\, y_0)$ belongs to the attracting basin $\mathcal{B}_0$. Since the right side of the inequality (A5) is a measurable function of x_0, defined and strictly positive almost everywhere, it follows that its integral is strictly positive. Evidently this integral is a lower bound for the area of $\mathcal{B}_0$. Thus $\mathcal{B}_0$ has positive measure as required.

The proof for the case $\mathrm{Lyap}(\mathcal{A}_0) > 0$ is completely analogous. However, it requires us to make use of the hypothesis that $f'_x(y)$ is strictly positive, even for $y = 0$, so that we can choose a small η with $0 < \eta < f'_x(0)$ everywhere, and with

$$\int \log\big(f'_x(0) - \eta\big)\, dx \;> 0\,. \tag{A6}$$

The estimate (A2) is then replaced by

$$f_x(y) \;\geq\; y\big(f'_x(0) - \eta\big) \qquad \text{whenever} \qquad y < \frac{\eta}{K}\,. \tag{A7}$$

Now suppose that the basin $\mathcal{B}_0$ has positive measure. Then, for a set of x_0 of positive measure, we could find orbits $(x_0,\, y_0) \mapsto (x_1,\, y_1) \mapsto \cdots$ which satisfied $0 < y_n < \frac{\eta}{K}$ for all n. But using (A6) and (A7) it is not hard to see that this is impossible. Therefore $\mathcal{B}_0$ has measure zero. The arguments for the basin $\mathcal{B}_1$ are completely analogous. □

Remark A.2. Lemma A.1 could also be seen as a consequence of the more general theory of Ergodic Attractors, see [Pugh and Shub, 1989, p. 4, Theorem 3]. For a beautiful survey of pioneering work in this area see [Barreira and Pesin, 2006, specially §§8.3 and 10.1].

Appendix B: Schwarzian Derivative and Cross-Ratios.

In this section, I, I', and I'' will denote intervals of real numbers. To any C^3-diffeomorphism $f : I \to I'$, there is associated the Schwarzian derivative $\mathcal{S}f : I \to \mathbb{R}$, as defined in Equation (3).

Notational Convention. We will write $\mathcal{S}f \prec 0$ to indicate that the inequality $\mathcal{S}f(y) < 0$ holds for y in a dense (and necessarily open) subset of I; and similarly $\mathcal{S}f \succ 0$ if $\mathcal{S}f(y) > 0$ in a dense open set.

PROPOSITION B.1. *The Schwarzian derivative has the following properties:*

1. *The sign of $\mathcal{S}f$ is preserved under composition: For example, given C^3-diffeomorphisms $I \overset{f}{\to} I' \overset{g}{\to} I''$ with $\mathcal{S}f \prec 0$ and $\mathcal{S}g \prec 0$, it follows that $\mathcal{S}(g \circ f) \prec 0$.*

2. *$\mathcal{S}f \prec 0$ if and only if $\mathcal{S}f^{-1} \succ 0$.*

3. *$\mathcal{S}f \prec 0$ if and only if the function $\varphi(y) = 1/\sqrt{|f'(y)|}$ is strictly convex (or in other words if and only if the function $\varphi'(y)$ is strictly increasing).*

4. *$\mathcal{S}f$ is identically zero if and only if f is a fractional linear transformation, $f(x) = (ax+b)/(cx+d)$, with $ad - bc \neq 0$ and with $cx + d \neq 0$ throughout I.*

Proof.

1. A straightforward calculation shows that the Schwarzian derivative of a composition is given by the formula
$$\mathcal{S}(g \circ f) \;=\; (f')^2\,\mathcal{S}g \,+\, \mathcal{S}f\,, \tag{B1}$$
and the conclusion follows easily. (It follows from this equation that $\mathcal{S}(g \circ f) = \mathcal{S}f$ if and only if $\mathcal{S}g = 0$.)

2. This follows by taking $g = f^{-1}$ in equation (B1) and noting that the identity map has Schwarzian zero.

3. It is not hard to calculate that the second derivative of $\varphi(y) = 1/\sqrt{|f'(y)|}$ satisfies the equation
$$\mathcal{S}f = -2\varphi''(x)/\varphi(x). \tag{B2}$$
Thus $\mathcal{S}f(y) < 0$ on a dense open set if and only if $\varphi''(y) > 0$ on a dense open set; and the assertion follows. (Similarly, $\mathcal{S}f \succ 0$ if and only if $\varphi''(y) < 0$ on a dense open set.)

4. From Equation (B2) we see that the Schwarzian is zero if and only if the function $\varphi(y)$ is linear, say $\varphi(y) = cy + d$ or in other words
$$f'(y) = \pm 1/(cy+d)^2.$$
Integrating, we see that this is true if and only if f is fractional linear. □

Definition. A fixed point $y = f(y) \in I \cap I'$ will be called *strictly attracting* if $|f'(y)| < 1$ and *strictly repelling* if $|f'(y)| > 1$.

LEMMA B.2. *A C^3-diffeomorphism $f : I \to I'$ with $\mathcal{S}f \prec 0$ can have at most three fixed points. If it has three fixed points, then the middle one must be strictly repelling and the other two must be strictly attracting.*

Proof. We must first show that there cannot be four distinct fixed points. If there were fixed points $y_0 < y_1 < y_2 < y_3$, then by the Mean Value Theorem each of the open intervals (y_0, y_1), (y_1, y_2), and (y_2, y_3) would contain a point at which the derivative $f'(y)$ is equal to $+1$. But this is impossible, since the function $\varphi(y) = 1/\sqrt{|f'(y)|}$ is strictly convex.

Now suppose that there are three fixed points $y_0 < y_1 < y_2$. Since $\varphi(y)$ is strictly convex, and takes the value $+1$ somewhere in (y_0, y_1) and also somewhere in (y_1, y_2). It follows easily that $\varphi(y_0) > 1 > \varphi(y_1) < 1 < \varphi(y_2)$, and the conclusion follows. □

Definition. The *cross-ratio* of four distinct real numbers will mean the expression
$$\rho(y_0, y_1, y_2, y_3) = \frac{(y_2 - y_0)(y_3 - y_1)}{(y_1 - y_0)(y_3 - y_2)}. \tag{B3}$$
(The reader should take care, since conflicting notations are often used.) Note that
$$\rho(y_0, y_1, y_2, y_3) > 1 \qquad \text{whenever} \qquad y_0 < y_1 < y_2 < y_3.$$
Evidently the cross ratio remains invariant whenever we replace each y_i by $ay_i + b$ with $a \neq 0$. A brief computation shows that it also remains invariant when we replace each y_i by $1/y_i$. Since every fractional linear transformation can be expressed as a composition of affine maps and inversions, it follows that the cross-ratio is invariant under fractional linear transformations.

We say that a monotone map f *increases cross-ratios* if
$$\rho\big(f(y_0), f(y_1), f(y_2), f(y_3)\big) > \rho(y_0, y_1, y_2, y_3) \qquad \text{whenever} \qquad y_0 < y_1 < y_2 < y_3.$$

LEMMA B.3. [ALLWRIGHT, 1978] *Again let $f : I \to I'$ be a C^3-diffeomorphism. Then f increases cross-ratios if and only if $\mathcal{S}f \prec 0$.*

Remark. In Lemmas B.2 and B.3, note that we can obtain a corresponding statement for the case $\mathcal{S}f \succ 0$ simply by applying the given statement to the inverse map from I' to I. For example: *f decreases cross-ratios if and only if $\mathcal{S}f \succ 0$.*

Proof of Lemma B.3. First suppose that $\mathcal{S}f < 0$ on a dense open set. Given points $y_0 < y_1 < y_2 < y_3$, after composing f with a fractional linear transformation, we may assume that f fixes the three points y_0, y_1, y_3. If $\mathcal{S}f \prec 0$, then y_0, y_3 are attracting and y_1 is repelling by Lemma B.2. Since there can be no fixed point between y_1 and y_3, it follows that f moves every intermediate point to the right. Thus $f(y_2) > y_2$, and it follows easily that f increases the cross-ratio $\rho(y_0, y_1, y_2, y_3)$.

Conversely, if $\mathcal{S}f$ is not negative on a dense open set, then it must either be strictly positive somewhere, or identically zero on some interval. In the first case, it would decrease some cross-ratio, and in the second case it would be fractional-linear and hence preserve cross-ratios within this interval. This completes the proof. □

Remark B.4. If $0 < y_1 < y_2 < 1$, then the *Poincaré distance* between y_1 and y_2 within $(0,1)$ can be defined as

$$d_{[0,1]}(y_1, y_2) \;=\; \int_{y_1}^{y_2} \Big(\frac{1}{y} + \frac{1}{1-y}\Big) dy \;=\; \log\ \rho(0, y_1, y_2, 1)\,.$$

This can be identified with the usual Poincaré distance within a complex disk having the interval $[0,1]$ as diameter. In terms of the *Poincaré arclength coordinate* of Equation (11), the Poincaré distance formula can also be written as,

$$d_{[0,1]}(y_1, y_2) = |t(y_2) - t(y_1)|. \tag{B4}$$

Now consider an orientation preserving diffeomorphism from a closed interval to itself.

LEMMA B.5. *If $\mathcal{S}f \prec 0$ for an orientation preserving diffeomorphism f of an interval $I = [a, b]$, then $f'(a)f'(b) < 1$.*

Proof. First consider the special case where $f'(a) = f'(b)$. If this common value were ≥ 1, then the convex function $\varphi(y)$ would be ≤ 1 at the two boundary points, and hence strictly less than one in the interior. It would follow that $f'(y) > 1$ for $a < y < b$. But this is impossible since I maps to itself.

For the general case, let $r : I \to I$ be the reflection which interchanges the two end points, and consider the auxiliary function $g = r \circ f \circ r$. Evidently $\mathcal{S}g \prec 0$, with $g'(a) = f'(b)$ and $f'(a) = g'(b)$. Hence the composition $f \circ g$ satisfies $\mathcal{S}(f \circ g) \prec 0$, and has derivative

$$(f \circ g)'(a) \;=\; (f \circ g)'(b) \;=\; f'(a)f'(b)\,.$$

The argument above then shows that this common value is strictly less than one, as required. □

Similarly $f'(a)f'(b) > 1$ if $\mathcal{S}f \succ 0$, and $f'(a)f'(b) = 1$ if $\mathcal{S}f \equiv 0$. These inequalities are clearly visible in Figure 2. The following characterization is closely related:

Assertion. *A* C^3*-diffeomorphism* $f : I \to I'$ *satisfies* $\mathcal{S}f \prec 0$ *if and only if, for any two points* $y_0 \neq y_1$ *in* I*, the product* $f'(y_0)f'(y_1)$ *is strictly less than the square of the slope* $\big(f(y_1) - f(y_0)\big)/(y_1 - y_0)$.

(There are corresponding characterizations for the case $\mathcal{S}f \succ 0$ or for $\mathcal{S}f \equiv 0$.) The proof, based on Lemma B.5, is not difficult.

References

Allwright, D. J.: *Hypergraphic functions and bifurcations in recurrence relations,* SIAM J. Appl. Math. **34** (1978), no.4, 687–691.

Barreira, L. and Y. B. Pesin: *Smooth Ergodic Theory and Nonuniformly Hyperbolic Dynamics,* In: "Handbook of Dynamical Systems," **1B**, Edited by B. Hasselblatt and A. Katok, Elsevier, Amsterdam, (2006), 57–264.

Bonifant, A.; Dabija, M.; and Milnor, J.: *Elliptic curves as attractors in* $\mathbb{P}^2$*,* IMS Preprint Series 2006-01, Arxiv math.DS/0601015, January 2006.

Erdös, P. and Kac, M.: *On the number of positive sums of independent random variables*, Bull. Amer. Math. Soc. **53** (1947), 1011–1020.

Feller, W.: "An Introduction to Probability Theory and its Applications," Wiley & Sons, Volume **1** (3rd edition) 1968.

———- : "An Introduction to Probability Theory and its Applications," Wiley & Sons, Volume **2** (2nd edition) 1966.

Kan, I.: *Open sets of diffeomorphisms having two attractors, each with an everywhere dense basin,* Bull. Amer. Math. Soc. **31** (1994), 68–74.

Kawada, Y. and Ito, K.: *On the probability distribution on a compact group. I.* Proc. Phys.-Math. Soc. Japan **22** (1940), 977–998.

Lévy, P.: *L'addition des variables aléatoires définies sur une circonférence*, Bull. Soc. Math. France **67** (1939), 1–40.

Pugh, S. and Shub, M.: *Ergodic Attractors,* Trans. of the Amer. Math. Soc. **312** (1989), no.1, 1–54.

ARACELI BONIFANT,
DEPARTMENT OF MATHEMATICS,
UNIVERSITY OF RHODE ISLAND,
KINGSTON, RI. 02881-0816
E-mail address: bonifant@math.uri.edu

JOHN MILNOR,
INSTITUTE FOR MATHEMATICAL SCIENCES,
STONY BROOK UNIVERSITY,
STONY BROOK, NY. 11794-3660.
E-mail address: jack@math.sunysb.edu