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 many-to-one skew product map of the cylinder $(\mathbb{R}/\mathbb{Z}) \times I$.

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.

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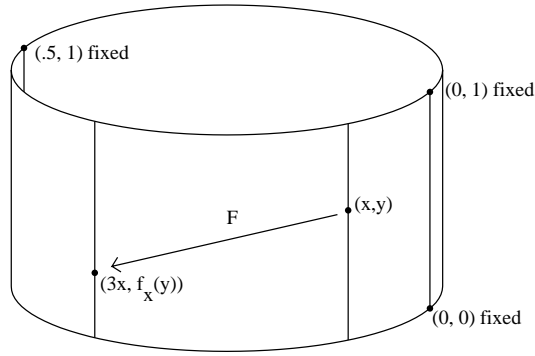
$$F(x,y) = \big(kx,\, f_x(y)\big), \tag{1}$$

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

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

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

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*
>
> $$\mathrm{Lyap}(\mathcal{A}_\iota) = \int\limits_{\mathbb{R}/\mathbb{Z}} \log\left(f_x'(\iota)\right) dx \qquad (2)$$
>
> *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

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

---

[1] 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 all orbits inside either lobe of the figure-eight spiral in towards an attracting equilibrium point. Then the figure-eight is a minimal measure attracting set with no dense orbit.

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 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.

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.3 (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 $\mathcal{B}_0 \cup \mathcal{B}_1$ has full measure. In fact, there is an almost everywhere defined measurable function $\sigma : \mathbb{R}/\mathbb{Z} \to I$ such that*

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

More generally, the same statement is true if the k-tupling map on the circle is replaced by any continuous ergodic transformation g on a compact space with g-invariant probability measure.

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 $0 \leq \sigma_0(x) \leq \sigma_1(x) \leq 1$ defined by the property that the orbit of (x, y):

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

---

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

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.4.) 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

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

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

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

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 $\beta(\mathcal{A}_\iota) = 0$.

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. $\qquad\square$

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.
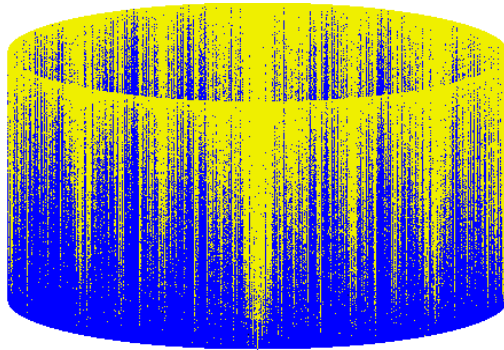

Fig. 2: Intermingled basins for the cylinder map F of Example 4.4.

Example 4.4. Following [KAN, 1994], let

$$q_a(y) = y \,+\, ay(1 - y)\,. \qquad (5)$$

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.4 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. 2.)

(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

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

  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, Theorem 3.2 asserts that the union of the attracting basins of the two boundaries $\Sigma \times \{0\}$ and $\Sigma \times \{1\}$ under the map $\widetilde{F}^{-1}$ 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.

---

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

- 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

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

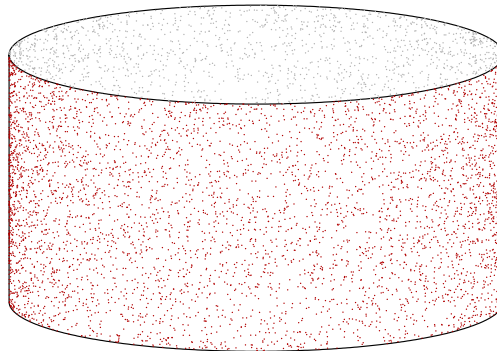This will be the required asymptotic measure for the original cylinder map F.

Fig. 3: *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 $p(x) \ne 0$. (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. 3). 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. Such a measure ν must necessarily be equal to the weak limit of $\big(\lambda + F_* \lambda + \cdots + F_*^{\circ n-1}\lambda\big)/n$ as $n \to \infty$; and 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. $\|\widetilde{x}\|/k \ \leq \ \|k\widetilde{x}\| \ \leq \ k\,\|\widetilde{x}\| \quad for\ all \ \ \widetilde{x} \in \Sigma.$

Proof. This follows easily from the definition. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

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. $\square$

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

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

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.4), 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

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

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 $s(\widetilde{x}) \leq 1$ 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 $r_0 < r(\sigma(\widetilde{x}),\, y) \leq r_1$. 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. $\qquad\square$

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.3 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, the same lemma implies that the set of orbits converging to $\mathcal{A}_0 \cup \mathcal{A}_1$ has measure zero, so that the Lemmas 5.4 through 5.6 apply.

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 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*

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

which varies over the entire real line for $0 < y < 1$, with inverse $y = e^t/(1 + e^t)$. If we embed the unit interval in the complex open disk of radius $1/2$ centered at $1/2$, then $|t|$ can be described as the distance from the midpoint, using the Poincaré metric for this disk. (Compare Appendix B.)

Since we are assuming $\mathcal{S}f_x$ identically zero, it follows that each f_x preserves cross-ratios or Poincaré distances. (See Equations (11) and (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 smooth function p from $\mathbb{R}/\mathbb{Z}$ to $\mathbb{R}$, we can set $c = p(x)$ to obtain an associated cylinder map

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

(The map F has an absolutely continuous invariant measure $dx\,dt$. However, this is not very useful since the total area $\int\!\int dx\,dt$ is infinite.)

Using the coordinate $t \in \mathbb{R}$ in place of $y \in (0,1)$, the cylinder map F will correspond to the map

$$(x,\, t) \;\mapsto\; \big(k\,x,\, t + p(x)\big)$$

of $(\mathbb{R}/\mathbb{Z}) \times \mathbb{R}$. The dynamics of F under iteration is governed by the average

$$A \;=\; \int_{\mathbb{R}/\mathbb{Z}} p(x)\, dx \tag{12}$$

of the displacement $p(x)$. For almost any orbit $(x_0, t_0) \mapsto \cdots \mapsto (x_n, t_n) \mapsto \cdots$, 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 A as $n \to \infty$. Thus if $A > 0$ 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 $A < 0$ then the lower boundary $\mathcal{A}_0$ will be a global attractor.

The borderline case where the average (12) is exactly zero, is much more interesting. We conjecture that the long term behavior of the sequence of numbers t_0, t_1, t_2, $\ldots$ is very much like that for a random walk, in which the successive differences $\Delta t_n = t_{n+1} - t_n$ are identically distributed independent random numbers with mean zero. In particular, we believe that the following theorem will be true whenever the periodic function $p(x)$ has average zero. However, the proof will apply only in the following very special case.

> **Hypothesis.** We now assume that $p(x)$ is a step function which takes a constant value on each of the k intervals $j/k \le x < (j+1)/k$.

If x_0 is randomly chosen, it then follows easily that the successive steps

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

actually are 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 entry in the base k expansion of x_0.) Choose some number $N \ge 0$ and define three sequences as follows. Let

  a_n be the number of integers $1 \le i \le n$ with $t_i > N$,
  b_n the number of such integers with $|t_i| \le N$, and
  c_n the number of such integers with $t_i < -N$.

Thus the associated frequencies a_n/n, b_n/n, c_n/n will have sum equal to $+1$.

THEOREM 6.1. *Let $p(x)$ be a step function as described above, not identically zero but with average $\int p(x)\,dx$ equal to zero. Then for arbitrary t_0 and for Lebesgue almost every x_0, the ratios b_n/n associated with the orbit $(x_0, t_0) \mapsto (x_1, t_1) \mapsto \cdots$ will converge to zero as $n \to \infty$, but a_n/n and c_n/n will not tend to any limit. In fact, these ratios vary so wildly that*

$$\liminf (a_n/n) \;=\; 0 \;<\; \limsup (a_n/n) \;=\; 1\,,$$

with a similar statement for c_n/n.

In terms of the original cylinder map, this means that most orbits spend most of the time extremely close to one or the other of the two boundaries; very occasionally jumping from one boundary to the other but in such an irregular way that there is no limiting asymptotic measure.

We are indebted to Harry Kesten and especially to Mikhail Lyubich for very substantial help with the proof. To begin the argument, note that for each t_0 and each $x_0 \in \mathbb{R}/\mathbb{Z}$ the number $L = \limsup_{n\to\infty}(a_n/n)$ is well defined, and is invariant under finite permutations of the sequence of differences Δt_n, or equivalently under finite permutations of the entries in the base k expansion of x_0. Therefore, according to the Hewitt-Savage *Zero-One Law*, each set

$$\{x_0 \;\; ; \;\; L \;\leq\; \text{constant}\}$$

has measure either zero or one. (See [FELLER **2**, 1966, IV.6].) It follows easily that L takes some constant value for almost all x_0.

First consider the case $N = 0$. We will prove that $L = 1$ except on a set of measure zero. Choose some small $\epsilon > 0$ and let $X_n(\epsilon)$ be the set of all x_0 for which $a_n/n > 1 - \epsilon$. Evidently the intersection

$$\bigcap_n \Big(X_n(\epsilon) \cup X_{n+1}(\epsilon) \cup X_{n+2}(\epsilon) \cup \cdots \Big)$$

is precisely the set of x_0 for which $L = \limsup(a_n/n)$ satisfies $L > 1 - \epsilon$. Thus, if we assume that $L \leq 1 - \epsilon$ on a set of x_0 of positive (and hence full) measure, then it would certainly follow that the measure of $X_n(\epsilon)$ must tend to zero as $n \to \infty$. But according to the *Arcsine Law*, the measure of $X_n(\epsilon)$ converges to the limit

$$1 - \frac{2}{\pi} \arcsin \sqrt{1-\epsilon} \;\;=\;\; \frac{2}{\pi} \arcsin \sqrt{\epsilon} \;\;>\;\; 0$$

as $n \to \infty$. (See [FELLER **1**, 1968, III.4; and **2**, 1966, XII.8].) This contradiction proves that $\limsup(a_n/n) = 1$ almost everywhere. Since $a_n + c_n \leq n$, it follows that $\liminf (c_n/n) = 0$; and a similar argument with a_n and c_n interchanged completes the proof for the case $N = 0$.

To complete the proof of Theorem 6.1, choosing some arbitrarily large N, we must show that the associated sequence b_n/n converges to zero for almost all x_0. The proof will be based on the following.

Consider a random walk $\tau_0, \tau_1, \ldots$ on the circle $\mathbb{R}/\mathbb{Z}$, starting with some specified τ_0, where the differences $\Delta\tau_n = \tau_{n+1} - \tau_n$ are identically distributed independent random variables which do not depend on τ_n.

LEMMA 6.2. *The resulting sequence $\{\tau_n\}$ is uniformly distributed around the circle with probability $+1$ if and only if the following condition is satisfied:*

(*) *There is no finite cyclic subgroup $G \subset \mathbb{R}/\mathbb{Z}$ such that $\Delta\tau_n \in G$ with probability $+1$.*

Proof. (Compare [Lévy, 1939], [Kawada and Ito, 1940].) Clearly this subgroup condition is necessary. To prove that it is sufficient, let μ be the common probability distribution for the differences $\Delta\tau_n$, let $S \subset \mathbb{R}/\mathbb{Z}$ be the support of μ, and let $S^{\mathbb{N}} = S \times S \times S \times \cdots$ be the space of sequences $\mathbf{s} = (s_0,\, s_1,\, \ldots)$ of elements of S, provided with the shift invariant measure $\mu^{\mathbb{N}} = \mu \times \mu \times \cdots$. It is not hard to see that the skew product map

$$F\bigl(\tau,\, s_0,\, s_1,\, \ldots\bigr) \;=\; (\tau + s_0,\, s_1,\, s_2,\, \ldots)$$

from $(\mathbb{R}/\mathbb{Z}) \times S^{\mathbb{N}}$ to itself preserves the measure $\lambda \times \mu^{\mathbb{N}}$, where λ is Lebesgue measure. We will prove that F is ergodic; or equivalently that:

Every bounded measurable F-invariant function $\phi : (\mathbb{R}/\mathbb{Z}) \times S^{\mathbb{N}} \;\to\; \mathbb{R}$ is constant almost everywhere.

Given such a function ϕ , let

$$\phi_n(\tau,\, s_0,\, \ldots,\, s_{n-1}) \;=\; \int \phi(\tau,\, \mathbf{s})\, d\mu^{\mathbb{N}}(s_n,\, s_{n+1},\, \ldots)$$

be the average of $\phi(\tau,\, \mathbf{s})$ over all possible choices of $s_n,\, s_{n+1},\, \ldots$. Clearly ϕ_n can also be expressed as an average over all choices of s_n,

$$\phi_n(\tau,\, s_0,\, \ldots,\, s_{n-1}) \;=\; \int \phi_{n+1}(\tau, s_0,\, \ldots,\, s_n)\, d\mu(s_n)\,.$$

On the other hand, using the condition $\phi = \phi \circ F$ of F-invariance, we see that $\phi_n(\tau,\, s_0,\, \ldots,\, s_{n-1})$ can also be described as the average of $\phi(\tau + s_0,\, s_1,\, s_2,\, \ldots)$ over all choices of $(s_n,\, s_{n+1},\, \ldots)$. But by definition, this is equal to $\phi_{n-1}(\tau + s_0,\, s_1,\, \ldots,\, s_{n-1})$. Thus, inductively, it follows that

$$\begin{aligned}\phi_n(\tau,\, s_0, \ldots,\, s_{n-1}) \;&=\; \phi_{n-1}(\tau + s_0,\, s_1,\, \ldots,\, s_{n-1}) = \phi_{n-2}(\tau + s_0 + s_1,\, s_2,\, \ldots,\, s_{n-1}) = \cdots \\ &=\; \phi_0(\tau + s_0 + \cdots + s_{n-1})\,.\end{aligned} \tag{13}$$

In particular, note that

$$\phi_0(\tau) \;=\; \int \phi_1(\tau,\, s)\, d\mu(s) \;=\; \int \phi_0(\tau + s)\, d\mu(s)\,. \tag{14}$$

If the condition **(*)** is satisfied, then we will use this last equation to prove that ϕ_0 is constant almost everywhere. Let

$$\widehat{\mu}(q) \;=\; \int \mathbf{e}(-qs)\, d\mu(s)$$

be the Fourier transform of μ, where $q \in \mathbb{Z}$, and where $\mathbf{e}(t)$ is an abbreviation for $e^{2\pi i t}$. Thus $\widehat{\mu}(q)$ is a weighted average of points on the unit circle; hence $|\widehat{\mu}(q)| \leq 1$. For $q \neq 0$, the condition **(*)** guarantees that the weight is not all concentrated at points s such that $qs \equiv 0 \pmod{\mathbb{Z}}$, so it follows that $\widehat{\mu}(q) \neq 1$.

The Fourier transform $\widehat{\phi}_0(q)$ is defined similarly as the integral of $\mathbf{e}(-q\tau)\phi_0(\tau)\,d\tau$. Recall that a bounded measurable function on the circle is constant almost everywhere if and only if its q-th Fourier coefficient is zero for every $q \neq 0$. Multiplying equation (14) by $\mathbf{e}(-q\tau)\,d\tau$ and then integrating, using the substitution $\eta = \tau + s$, we obtain

$$\widehat{\phi}_0(q) \;=\; \int \mathbf{e}(-q\eta)\,\phi_0(\eta)\,d\eta \int \mathbf{e}(qs)\,d\mu(s) \;=\; \widehat{\phi}_0(q)\,\widehat{\mu}(-q)\,.$$

For every $q \neq 0$, since $\widehat{\mu}(-q) \neq 1$, it follows that $\widehat{\phi}_0(q) = 0$. Therefore ϕ_0 takes some constant value v almost everywhere. Using equation (13), it follows that every ϕ_n also takes the value v almost everywhere.

We can now prove that ϕ is constant almost everywhere. Suppose to the contrary, for example, that $\phi(\tau, \mathbf{s}) > v + \epsilon > 0$ on a set Σ of positive measure. Choose a point of density $(\tau^*, \mathbf{s}^*)$ for Σ. Since ϕ is bounded, it would follow that the average of ϕ over a small neighborhood of $(\tau^*, \mathbf{s}^*)$ is strictly greater than v. But the fact that each ϕ_n equals v almost everywhere implies that every such average is also equal to v. Thus ϕ must be constant almost everywhere; which proves ergodicity.

Now the Birkhoff Ergodic Theorem implies that almost every orbit of F is uniformly distributed with respect to the measure $\lambda \times \mu^{\mathbb{N}}$. Projecting to the first coordinate, it follows that almost every sequence

$$\tau, \quad \tau + s_0, \quad \tau + s_0 + s_1, \quad \ldots$$

is uniformly distributed with respect to λ. This completes the proof of Lemma 6.2. $\qquad\square$

Proof of Theorem 6.1 (conclusion). Recall that b_n is the number of $0 \leq j < n$ with $|t_j| \leq N$. We must show that the ratio b_n/n tends to zero with probability $+1$ as $n \to \infty$. Choosing some large number $L \gg N$, the quotients $t_n/L \pmod{\mathbb{Z}}$ form a random walk on the circle. If L is chosen so that the ratio $\Delta t_n/L$ is irrational with probability > 0, then by Lemma 6.2, these quotients are uniformly distributed around the circle. Let $b_n(L) \geq b_n$ be the number of $j \in [0, n)$ with t_j congruent to an element of $[-N, N]$ modulo $L\mathbb{Z}$. Then it follows that $b_n(L)/n$ converges to $2N/L$ as $n \to \infty$. Since L can be arbitrarily large, this proves that $b_n/n \to 0$, as required. $\qquad\square$

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

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

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

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

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

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

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

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

is certainly defined and finite for almost all x. Now suppose that

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

Then a straightforward induction shows that

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

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. $\qquad\square$

Appendix B: Schwarzian Derivative and Cross-Ratios.

Recall that the Schwarzian derivative $\mathcal{S}f$ of a C^3 interval diffeomorphism $f : I \to I$ was defined in Equation (3). The statement $\mathcal{S}f < 0$ will mean that the inequality $\mathcal{S}f(y) < 0$ holds for y in a dense open subset of I; and similarly for $\mathcal{S}f > 0$ (or $\mathcal{S}f = 0$).

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

1. *The sign of $\mathcal{S}f$ is preserved under iteration of f. For example if $\mathcal{S}f < 0$ and $\mathcal{S}g < 0$, then $\mathcal{S}(f \circ g) < 0$.*

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

3. *$\mathcal{S}f < 0$ if and only if the function $\varphi(y) = 1/\sqrt{|f'(y)|}$ is concave upwards.*

4. *$\mathcal{S}f = 0$ if and only if f is a fractional linear transformation, $f(x) = (ax+b)/(cx+d)$ where we may assume that $ad - bc = \pm 1$.*

Proof.

1. A straightforward calculation shows that the Schwarzian derivative of a composition is given by the formula

$$\mathcal{S}(f \circ g) \;=\; (g')^2 \, \mathcal{S}f \,+\, \mathcal{S}g \,, \tag{B1}$$

and the conclusion follows easily.

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.

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. $\square$

Now consider a C^3-diffeomorphism $f \,:\, I \,\to\, I'$ where I' may be a different interval of real numbers. 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. *Suppose that $\mathcal{S}f < 0$ throughout a dense open subset of I. If f has two fixed points, then at least one must be strictly attracting. Furthermore, if there are three fixed points then the middle one must be strictly repelling and the other two must be strictly attracting. Such a map can never have four fixed points.*

Proof. Given two fixed points $\alpha < \beta$, we will first show that at least one of the two must be strictly attracting. In fact, it follows from the Mean Value Theorem that some point $\alpha < x < \beta$ must satisfy $f'(y) = 1$. If both $f'(\alpha) \geq 1$ and $f'(\beta) \geq 1$, then the graph of f' would have to have a local minimum somewhere in the open interval (α, β). Hence the function $\varphi(y)$ would have a local maximum. Since $\varphi''(y) > 0$ on a dense open set, this clearly leads to a contradiction.

Now consider three fixed points $\alpha < \beta < \gamma$. Then The Mean Value Theorem yields points $y \in (\alpha, \beta)$ and $y' \in (\beta, \gamma)$ with $f'(y) = f'(y') = 1$. Again, if $f'(\beta) \leq 1$, then the graph of f' would have a local minimum, yielding a contradiction. Therefore β is strongly repelling, hence α and γ must be strongly attracting. Evidently this leaves no possibility for a fourth fixed point. $\square$

LEMMA B.3. *If $\mathcal{S}f < 0$ for an orientation preserving diffeomorphism $f : I \to I$, then $f'(0)f'(1) < 1$.*

Proof. By the previous lemma, at least one of the two boundary fixed points must be strictly attracting: either $f'(0) < 1$, or $f'(1) < 1$, or both. Now consider the auxiliary function

$$g(y) \ = \ 1 - f(1 - y) \,.$$

Evidently g also has negative Schwarzian, with $g'(0) = f'(1)$ and $f'(0) = g'(1)$. The composition $f \circ g$ also has negative Schwarzian; hence $f \circ g$ has derivative less than one at at least one of the two endpoints. But the derivative at either endpoint is equal to the product $f'(0) f'(1)$. It follows that this product is less than one, as required.   □

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 $a y_i + b$ 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 will 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.4. (Allwright.) *Again let $f : I \to I'$ be a C^3-diffeomorphism. Then f increases cross-ratios if and only if $\mathcal{S}f < 0$ throughout some dense open subset of I.*

Remark. In Lemmas B.2, B.3, and B.4, note that we can obtain a corresponding statement for the case $\mathcal{S}f > 0$ simply by applying the given statement to the inverse map from I' to I. For example: *f decreases cross-ratios if and only if $\mathcal{S}f > 0$ on a dense open set.*

Proof of Lemma B.4. (Compare [Allwright, 1978].) 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 $\mathcal{S}f < 0$ on a dense set, 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.5. If $0 < y_1 < y_2 < 1$, then the *Poincaré distance* between y_1 and y_2 within $(0,1)$ can be defined as

$$
\begin{aligned}
d_{[0,1]}(y_1, y_2) &= \int_{y_1}^{y_2} \left(\frac{1}{y} + \frac{1}{(1-y)} \right) dy \\
&= \log \left(\frac{y_2(1-y_1)}{y_1(1-y_2)} \right) = \log \rho(0, y_1, y_2, 1).
\end{aligned}
$$

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,

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

References

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

[BONIFANT, DABIJA AND MILNOR, 2006] 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.

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

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

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

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

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

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