Ergod. Th. & Dynam. Sys. (2001), **21**, 1843–1854
DOI: 10.1017/S0143385701001882

Minimal non-ergodic C^1-diffeomorphisms of the circle

FERNANDO OLIVEIRA† and L. F. C. DA ROCHA‡

† *Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, Av. Antonio Carlos, 6627, CEP 30123-970, Belo Horizonte/MG, Brazil (e-mail: fernando@mat.ufmg.br)*

‡ *Instituto de Matemática, Universidade Federal do Rio Grande do Sul, Av. Bento Goncalves 9500, CEP 91500, Porto Alegre/RS, Brazil (e-mail: rocha@mat.ufrgs.br)*

(*Received* 21 *December* 1998 *and accepted in revised form* 2 *November* 2000)

Abstract. We construct, for each irrational number α, a minimal C^1-diffeomorphism of the circle with rotation number α which is not ergodic with respect to the Lebesgue measure.

1. *Introduction*

A diffeomorphism f of the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is ergodic with respect to the Lebesgue measure if there is no f-invariant Borel set with Lebesgue measure strictly between zero and one. In [**1**], Denjoy proved that a C^1-diffeomorphism with bounded variation derivative is ergodic. In the other direction, Denjoy constructed examples of C^1-diffeomorphism in any rotation class with invariant Cantor sets of positive measure. These examples, having wandering intervals, are not minimal. In this paper we construct, for each irrational number $\alpha \in (0, 1)$, an orientation preserving C^1-diffeomorphism of the circle which is minimal, i.e. has every orbit dense, but is not ergodic.

Given an orientation preserving C^1-diffeomorphism $f : \mathbb{T} \to \mathbb{T}$, to define its rotation number, $\rho(f)$, lift f to $\widetilde{f} : \mathbb{R} \to \mathbb{R}$ and take

$$\rho(f) := \lim_{n\to\infty} \frac{\widetilde{f}^n(x) - x}{n}$$

for $x \in \mathbb{R}$, where $\widetilde{f}^n$, $n \in \mathbb{Z}$, denotes the iterates of $\widetilde{f}$. An equivalent, more enlightening, combinatorial definition can be found in de Melo and van Strien [**4**, p. 33]. If $\rho(f) = \alpha$ and f is minimal then f is conjugated to the rotation $R_\alpha : x \in \mathbb{T} \longmapsto x + \alpha \in \mathbb{T}$. This means that there is an orientation-preserving homeomorphism $h : \mathbb{T} \to \mathbb{T}$ such that $h \circ R_\alpha = f \circ h$.

To construct f, we go the other way around; we construct an homeomorphism h such that $f := h \circ R_\alpha \circ h^{-1}$ is a C^1-diffeomorphism with the required properties. We define h

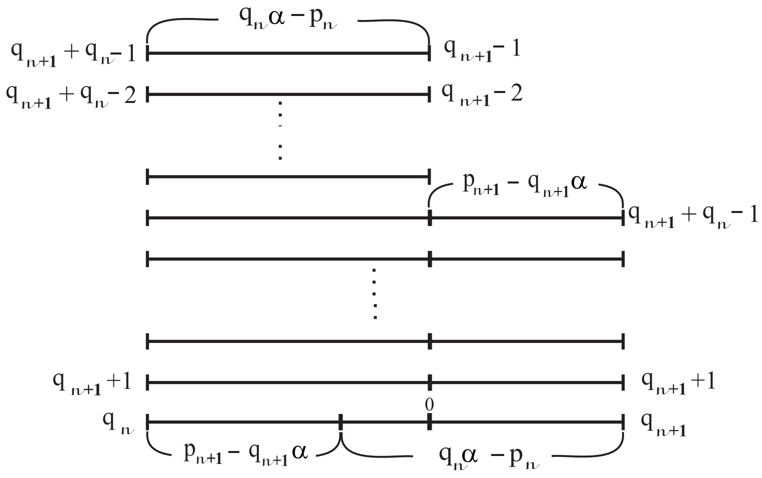

FIGURE 1.

first as an order-preserving map on $\mathcal{O}$, the non-negative orbit of $0 = 1$ under R_α, in such a way that $h(\mathcal{O})$ is dense in $\mathbb{T}$. Then, it is immediate, h extends to $\mathbb{T}$ as an homeomorphism. To define h on $\mathcal{O}$ we need to understand the relationship between the dynamical and linear orders on $\mathcal{O}$, an information which is encoded in the continued fraction expansion of α. We recall this classical formalism in a way that fits our needs in §2. In §3 we collect the lemmas we will need to prove, in §4, the following theorem.

THEOREM 1. *Given an irrational number α, $0 < \alpha < 1$, there is a minimal orientation-preserving C^1-diffeomorphism, $f : \mathbb{T} \longrightarrow \mathbb{T}$, which is not ergodic with respect to Lebesgue measure and such that $\rho(f) = \alpha$.*

2. *Continued fractions and towers*

Fix $\alpha \in (0, 1)$ as irrational and let $a_1, a_2, a_3, \ldots$ be the sequence of its partial quotients and p_n/q_n be the sequence of its approximants. Let $\{\mathcal{T}\}$, $\mathcal{T} = \mathcal{T}(aq_{n+1} + q_n)$, be the set of partitions of $\mathbb{T}$ by intervals with extremes

$$\{R_\alpha^i(0)\}_{i=0}^{aq_{n+1}+q_n-1},$$

for $n \geq 1$ and $1 \leq a \leq a_{n+2}$, well ordered by the relation of refinement.

For easy reference it is convenient to stack the intervals of $\mathcal{T}$ into a pair of towers so that we get from $\mathcal{T}$ to its next refinement, $\widetilde{\mathcal{T}}$, by the process of ‘cutting and stacking’. This procedure is explained in detail in [**3**, Appendix 1] (actually, Katznelson and Ornstein work with ‘towers with balconies’ separating balconies from towers give our towers). Figure 1 shows, $\mathcal{T}(q_{n+1} + q_n)$ for n even. In this figure the integer k at the side of a level stands for $R_\alpha^k(0)$. In general, $\mathcal{T}(aq_{n+1} + q_n)$ has a left tower made of left intervals or L-intervals

$$\{R_\alpha^i[R_\alpha^{(a-1)q_{n+1}+q_n}(0), 0]\}_{i=0}^{q_{n+1}-1}$$

and a right tower made of R-intervals

$$\{R_\alpha^i[0, R_\alpha^{q_{n+1}}(0)]\}_{i=0}^{(a-1)q_{n+1}+q_n-1}$$

for n odd; for n even change left for right.

The idea for constructing h, and therefore f, is very simple: at each stage in the process of forming towers, points of $\mathbb{T}$ fall either in a L- or R-interval. Suppose we choose a side, say right, for a fixed increasing subsequence $\mathcal{T}_k$ of the above sequence of towers, which, it should be emphasized, is entirely determined by (the continued fraction expansion of) α. The set of points S which are not in the orbit of zero but which are in an interval of the right side of $\mathcal{T}_k$ for k arbitrarily large is clearly R_α-invariant. We are going to construct h in such a way that $1 > \mu(h(S)) > 0$ where μ is the Lebesgue measure. This is easily done. We just have to imitate the construction of the Cantor map and conveniently distort the intervals in the right side as they appear in the process of refinement. If we do that, then $f := h \circ R_\alpha \circ h^{-1}$ is a homeomorphism but, of course, not, in general, a C^1-diffeomorphism. Our task is then to show that we can accomplish this change of lengths and also get a diffeomorphism. To control those distortions we will need the following lemma.

LEMMA 2. *Let $\mathcal{T}$ be a tower of R_α. Denote the L-intervals of $\mathcal{T}$ by L_i and the R-intervals by R_i. Cut and stack $\mathcal{T}$ to get a new tower $\widetilde{\mathcal{T}}$. Denote the intervals in the left tower of $\widetilde{\mathcal{T}}$ by l_j and similarly define the intervals r_j. These smaller intervals decompose each left interval of $\mathcal{T}$ into n^l and m^l intervals, respectively. Analogously they decompose the right intervals of $\mathcal{T}$ into n^r and m^r intervals, respectively. Then, cutting and stacking, we get infinitely often pairs of towers $\widetilde{\mathcal{T}}$ such that*

$$\frac{1}{5} < \frac{m^l r}{n^l l} < \rho := \frac{(m^l + m^r)r}{(n^l + n^r)l} < \frac{m^r r}{n^r l} < 5.$$

Proof. Since our thesis is invariant under scaling, we can consider the rescaled first return map to the bottom of the towers and assume the initial towers with height one. Cut and stack $\mathcal{T}$ to get $\widetilde{\mathcal{T}}$ with n^l, n^r, m^l and m^r. Then $m^r + m^l$ is the height of the right tower of $\widetilde{\mathcal{T}}$ and $n^r + n^l$ the height of its left tower. We can take these quantities as large as we please since, say, $n^r/n^r + n^l$ goes to α as we cut and stack, by the unique ergodicity of R_α. We have

$$\begin{pmatrix} L \\ R \end{pmatrix} = \begin{pmatrix} n^l & m^l \\ n^r & m^r \end{pmatrix} \begin{pmatrix} l \\ r \end{pmatrix}$$

and $n^l m^r - n^r m^l = 1$ since this matrix is a product of the matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

we have $m^l r/n^l l < m^r r/n^r l$. We will show that by further cutting and stacking of $\widetilde{\mathcal{T}}$ we can get $\widetilde{\widetilde{\mathcal{T}}}$ such that $\frac{1}{2} \leq \rho \leq 4$, δ^- and $\delta^+ \leq \frac{1}{4}$, where

$$\delta^- = \rho - \frac{m^l r}{n^l l} = \frac{1}{n^l(n^r + n^l)} \frac{r}{l} \quad \text{and} \quad \delta^+ = \frac{m^r r}{n^r l} - \rho = \frac{1}{n^r(n^r + n^l)} \frac{r}{l}$$

which will prove the lemma. Consider two cases: either $a_n = 1$ for n large enough or $a_n > 1$ for n arbitrarily large. Since our thesis is invariant by flipping sides, left and right, we can suppose, without loss of generality, that $r > l$ and $n^r + n^l < m^r + m^l$. If the second possibility holds, cut and stack $\mathcal{T}$ to get $\widetilde{\mathcal{T}}$ critical immediately preceding the appearance of a partial quotient $a_n = [r/l] > 1$. Cut and stack $\widetilde{\mathcal{T}}$ $a := [a_n/2] \geq 1$ times to get $\widetilde{\widetilde{\mathcal{T}}}$. We have

$$L = n^l l + m^l r = (n^l + m^l a)l + m^l(r - al)$$
$$R = n^r l + m^r r = (n^r + m^r a)l + m^r(r - al)$$

and therefore

$$\rho = \frac{(m^r + m^l)(r - al)}{((n^r + m^r a) + (n^l + m^l a))l} = \frac{a_n + f - a}{(n^r + n^l)/(m^r + m^l) + a}$$

where f is the fractional part of r/l. Thus

$$\frac{a_n - a}{1 + a} \leq \rho \leq \frac{a_n - a + 1}{a}$$

and, using the definition of a, we have

$$\frac{1}{2} \leq \frac{a}{1 + a} \leq \rho \leq \frac{a + 3}{a} \leq 4$$

as required. For δ^+ we have

$$\begin{aligned}\delta^+ &= \frac{1}{(n^r + m^r a)(n^r + m^r a + n^l + m^l a)} \frac{r - al}{l} \\ &= \frac{1}{(n^r + m^r a)(m^r + m^l)} \frac{(r/l) - a}{(n^r + n^l)/(m^r + m^l) + a} \\ &\leq \frac{4}{m^r + m^l} \leq \frac{1}{4}\end{aligned}$$

and similar estimates hold for δ^-.

Now suppose $a_n = 1$ for n large enough. Cut and stack $\mathcal{T}$ to get $\widetilde{\mathcal{T}}$ as the partial quotients a_n converge to 1. Cut and stack $\widetilde{\mathcal{T}}$ twice to get $\widetilde{\widetilde{\mathcal{T}}}$. We have

$$L = (n^l + m^l)(2l - r) + (2m^l + n^l)(r - l)$$
$$R = (n^r + m^r)(2l - r) + (2m^r + n^r)(r - l)$$

and therefore

$$\rho = \frac{(2m^r + n^r + 2m^l + n^l)(r - l)}{(n^r + m^r + n^l + m^l)(2l - r)} = \left(1 + \frac{m^r + m^l}{n^r + m^r + n^l + m^l}\right) \frac{(r - l)}{(2l - r)}$$

which shows that $1 \leq \rho \leq 4$ as required, since $0 \leq r - l < l$, $0 \leq l - (r - l) = 2l - r < r - l$ and $0 \leq (r - l) - (2l - r) < 2l - r$. As for, say, δ^+ we have, using the above inequalities, that

$$\delta^+ = \frac{1}{(n^r + m^r)(n^r + m^r + n^l + m^l)} \frac{r - l}{2l - r} < \frac{1}{4}. \qquad \square$$

3. *Pl towers*

Now take a pair of towers $\mathcal{T}$ of R_α and change the lengths of their intervals except for the top and bottom ones. This change is subject only to the condition that the sum of the lengths of all intervals is one. The new pair of towers, $\widetilde{\mathcal{T}}$, with its scheme of mappings and identifications, defines, in the obvious way, a unique (modulo rotations) piecewise (pl) homeomorphism of the circle, f, and a pl conjugacy, h, between f and R_α. We just have to take every map in sight as orientation preserving and affine. More precisely we consider towers, which we shall call pl towers, such that:

(1) the combinatorics of the towers are the combinatorics of some R_α tower;

(2) the maps one floor up are affine;

(3) the top and bottom levels are the same as in the corresponding R_α tower and therefore the top to bottom maps are the same isometries as for R_α.

Since an orientation-preserving affine homeomorphism between intervals is unique, any diagram of such maps commutes and we have a pl conjugacy, h, between f and R_α on the mid levels. As the top and bottom levels were unchanged the same holds there and the conjugacy thus defined ensures that $\rho(f) = \alpha$ and the minimality of f. Therefore h is affine on all levels and, in fact, the identity on the top and bottom levels, the iterates of 0 under f coincide with the corresponding ones of R_α on the top and bottom levels and are an affine image of them on the other levels.

Observe that if we have one of these towers and we cut and stack its intervals (as prescribed by the pl homeomorphism f it defines) any number of times and change their lengths in the allowed way, we still get a tower of the same sort which defines a pl homeomorphism f with rotation number α. Also, if we take one of these towers and start to cut and stack, the maximum length of the intervals in the towers goes down monotonically to zero since the cutting orbit is dense in $\mathbb{T}$.

The derivative of a pl homeomorphism defines a positive step map φ which, by definition, is a real-valued map defined on $\mathbb{T}$ such that there are distinct points $p_0, p_1, \ldots, p_k = p_0 \in \mathbb{T}$, indexed in the counterclockwise sense, such that φ is constant in each interval $[p_{i-1}, p_i)$. If p is a point in $\mathbb{T}$ we define the jump of φ at p as

$$J(\varphi, p) = \left| \lim_{x \to p^+} \varphi(x) - \lim_{x \to p^-} \varphi(x) \right|.$$

If φ_n is a sequence of step maps, satisfying:

(i) $\sum^n \|\varphi_{n+1} - \varphi_n\| < \infty$ where $\|\varphi\|$ is the supremum norm of φ;

(ii) $\max_{p \in \mathbb{T}^1} J(\varphi_n, p) \to 0$ as $n \to \infty$;

then clearly φ_n converges uniformly to a continuous map φ. Integrating we get the following lemma.

LEMMA 3. *Let f_n be a sequence of pl homeomorphisms of $\mathbb{T}$ converging pointwise to a pl homeomorphism f. Suppose the sequence of derivatives $\varphi_n = Df_n$ satisfies (i) and (ii) above and is uniformly bounded away from zero. Then f is a C^1-diffeomorphism and $Df = \varphi$ where $\varphi = \lim \varphi_n$.*

Let I_1 and I_2 be two intervals, and $f : I_1 \to I_2$ be the affine orientation-preserving homeomorphism between them. Let $\mathcal{T}_1 = \{p_{1j}\}_{j=0}^k$ be a partition of I_1 and

$\mathcal{T}_2 = f(\mathcal{T}_1) = \{p_{2j}\}$. Fix two indices $1 < k_1 < k_2 < k$ and move the points of $\mathcal{T}_1$ without changing their relative order to form a new partition $\widetilde{\mathcal{T}}_1 = \{\widetilde{p}_{1j}\}$ of I_1 and do the same for $\mathcal{T}_2$ thus getting $\widetilde{\mathcal{T}}_2 = \{\widetilde{p}_{2j}\}$. We are no longer assuming that $\widetilde{\mathcal{T}}_2 = f(\widetilde{\mathcal{T}}_1)$. Consider $\widetilde{f} : I_1 \to I_2$ as the orientation-preserving pl homeomorphism defined by the partitions $\widetilde{\mathcal{T}}_1$ and $\widetilde{\mathcal{T}}_2$. Suppose the change is such that:

(1) $\widetilde{p}_{ik_1} = p_{ik_1}$ and $\widetilde{p}_{ik_2} = p_{ik_2}$, $i = 1, 2$, i.e. the marked points are unchanged;

(2) $\widetilde{f}$ is affine in the intervals $[\widetilde{p}_{11}, \widetilde{p}_{1k_1}]$ and $[\widetilde{p}_{1k_2}, \widetilde{p}_{1k-1}]$;

(3) the intervals of $\mathcal{T}_1$ (respectively $\mathcal{T}_2$) in $[\widetilde{p}_{1k_1}, \widetilde{p}_{1k_2}]$ (respectively $[\widetilde{p}_{2k_1}, \widetilde{p}_{2k_2}]$) are partitioned in three groups of intervals $\mathcal{C}_1$, $\mathcal{E}_1$ and $\mathcal{Q}_1$ (their image by f being respectively $\mathcal{C}_2$, $\mathcal{E}_2$ and $\mathcal{Q}_2$). Denoting the corresponding intervals of $\widetilde{\mathcal{T}}_1$ (respectively $\widetilde{\mathcal{T}}_2$) by adding $\widetilde{\ }$ we assume that the intervals of $\widetilde{\mathcal{C}}_1$ (respectively $\widetilde{\mathcal{C}}_2$) are contracted by a factor of λ_1, $0 < \lambda_1 < 1$ (respectively λ_2, $0 < \lambda_2 < 1$). The intervals of $\widetilde{\mathcal{Q}}_1$ (respectively $\widetilde{\mathcal{Q}}_2$) remain with length equal to their counterparts in $\mathcal{T}_1$ (respectively $\mathcal{T}_2$). The remaining intervals in $\widetilde{\mathcal{E}}_1$ (respectively $\widetilde{\mathcal{E}}_2$) are, consequently, expanded. We assume this expansion is uniform.

LEMMA 4. *Under the above hypothesis, defining* $c = \mu(\cup\mathcal{C}_1)$, $e = \mu(\cup\mathcal{E}_1)$, $s = \mu(\cup\mathcal{Q}_1)$, $\kappa = \max\{c/e, e/c\}$ *and*

$$r = \max\left\{ \frac{\widetilde{p}_{i1} - p_{i0}}{p_{ik_1} - p_{i0}}, \frac{p_{ik} - \widetilde{p}_{ik-1}}{p_{ik} - p_{ik_2}} \;\middle|\; i = 1, 2 \right\},$$

and assuming $r < 1/2$, *we have*

$$\|D\widetilde{f} - Df\| \le \begin{cases} \kappa|\lambda_2/\lambda_1 - 1|\|Df\|, & \text{for intervals in } [\widetilde{p}_{1k_1}, \widetilde{p}_{1,k_2}] \\ 4r\|Df\|, & \text{for intervals in } [\widetilde{p}_{11}, \widetilde{p}_{1k_1}] \text{ or } [\widetilde{p}_{1k_2}, \widetilde{p}_{1k-1}]. \end{cases}$$

Proof. For an interval $[\widetilde{p}_{1j-1}, \widetilde{p}_{1j}]$ in $\widetilde{\mathcal{C}}_1$ we have

$$\|D\widetilde{f} - Df\| = \left|\frac{\lambda_2}{\lambda_1} - 1\right| \frac{p_{2j} - p_{2j-1}}{p_{1j} - p_{1j-1}} \le \left|\frac{\lambda_2}{\lambda_1} - 1\right| \|Df\|.$$

If $[\widetilde{p}_{1j-1}, \widetilde{p}_{1j}]$ is in $\widetilde{\mathcal{E}}_1$ we have in the same way

$$\|D\widetilde{f} - Df\| \le \left|\frac{\nu_2}{\nu_1} - 1\right| \|Df\|$$

where ν_i is the expansion undergone by the intervals in $\widetilde{\mathcal{E}}_i$, $i = 1, 2$. Now $c + e + s = p_{1k_2} - p_{1k_1} = \lambda_1 c + \nu_1 e + s$ and $a(p_{1k_2} - p_{1k_1}) = p_{2k_2} - p_{2k_1} = \lambda_2 ac + \nu_2 ae + as$, where $a = |I_2|/|I_1|$. Then $c + e = \lambda_1 c + \nu_1 e = \lambda_2 c + \nu_2 e$ from which we get

$$\begin{aligned} \left|\frac{\nu_2}{\nu_1} - 1\right| &= \left|\frac{[(c+e)/c] - \lambda_2}{[(c+e)/c] - \lambda_1} - 1\right| = \frac{|\lambda_1 - \lambda_2|}{[(c+e)/c] - \lambda_1} \\ &\le \frac{c}{e}\left|\frac{\lambda_2}{\lambda_1} - 1\right| \le \kappa\left|\frac{\lambda_2}{\lambda_1} - 1\right|. \end{aligned}$$

For intervals in $[\widetilde{p}_{11}, \widetilde{p}_{1k_1}]$ (or $[\widetilde{p}_{1k_2}, \widetilde{p}_{1k-1}]$) we have

$$\begin{aligned}\|D\widetilde{f} - Df\| &= \left|\frac{p_{2k_1} - \widetilde{p}_{21}}{p_{1k_1} - \widetilde{p}_{11}} - \frac{p_{2k_1} - p_{20}}{p_{1k_1} - p_{10}}\right| \\ &= \left|\frac{p_{2k_1} - p_{20} - (\widetilde{p}_{21} - p_{20})}{p_{1k_1} - p_{10} - (\widetilde{p}_{11} - p_{10})} - \frac{p_{2k_1} - p_{20}}{p_{1k_1} - p_{10}}\right| \\ &\leq 2\frac{(p_{1k_1} - p_{10})(\widetilde{p}_{21} - p_{20}) + (p_{2k_1} - p_{20})(\widetilde{p}_{11} - p_{10})}{(p_{1k_1} - p_{10})^2} \\ &\leq 2\left(\frac{\widetilde{p}_{21} - p_{20}}{p_{2k_1} - p_{20}} + \frac{\widetilde{p}_{11} - p_{10}}{p_{1k_1} - p_{10}}\right)\frac{p_{2k_1} - p_{20}}{p_{1k_1} - p_{10}} \leq 4r\|Df\|. \qquad \square\end{aligned}$$

Using that Df is a step map and the triangle inequality we have

$$|JD\widetilde{f}(\widetilde{p})| \leq \begin{cases} 2\kappa|\lambda_2/\lambda_1 - 1|\|Df\|, & \text{for } \widetilde{p} \in (\widetilde{p}_{1k_1}, \widetilde{p}_{1k_2}), \\ (\kappa|\lambda_2/\lambda_1 - 1| + 4r)\|Df\|, & \text{for } \widetilde{p} = \widetilde{p}_{1k_1} \text{ or } \widetilde{p} = \widetilde{p}_{1k_2}, \\ |D\widetilde{f}(p_{10}) - Df(p_{10})| + 4r\|Df\|, & \text{for } \widetilde{p} = \widetilde{p}_{11}, \\ |D\widetilde{f}(p_{1k}) - Df(p_{1k})| + 4r\|Df\|, & \text{for } \widetilde{p} = \widetilde{p}_{1k-1}. \end{cases}$$

Since the hypothesis is symmetric we have similar estimates for $D\widetilde{f}^{-1}$ and $JD\widetilde{f}^{-1}$, changing λ_2/λ_1 to λ_1/λ_2.

Let $\mathcal{T} = \mathcal{T}(aq_{n+1} + q_n)$ be a pl tower defined by a pl homeomorphism f, $\rho(f) = \alpha$. Denote its intervals by $\{L_i\}_{i=0}^{h^l-1}$ and $\{R_i\}_{i=0}^{h^r-1}$, where the heights are, say, $h^l = aq_{n+1} + q_n$ and $h^r = q_{n+1}$, indexed from bottom to top, and define $S = \cup R_i$. Fix positive integers $0 < h < \min\{h^l, h^r\}$, $\widetilde{h}$ and $q^r \in R_0$, $q^l \in L_0$ so close to zero that $R_\alpha^{-1}(0) \in f^{h^r-1}[q^r, f^{h^l}(q^l)] \cup f^{h^l-1}[f^{h^r}(q^r), q^l]$ but not in the R_α-orbit of zero.

Cut and stack $\mathcal{T}$ to a tower $\widetilde{\mathcal{T}} = \widetilde{\mathcal{T}}(\widetilde{a}q_{\widetilde{n}+1} + q_{\widetilde{n}})$ and define l- and r-intervals, n^l, m^l, n^r and m^r as in Lemma 2. We suppose $0 < \widetilde{h} < \widetilde{h}^l/2$ and $\widetilde{h}^r/2$, the heights of $\widetilde{\mathcal{T}}$, where we have started to use the convention of adding $\widetilde{\ }$ to the names of objects referring to $\widetilde{\mathcal{T}}$. Define $\mathcal{X} = \mathcal{X}(h)$ as the set of intervals on the first or top h levels of $\mathcal{T}$ and $\widetilde{\mathcal{D}} = \widetilde{\mathcal{D}}(q^r, q^l)$ as the complement of the set of intervals in $\widetilde{\mathcal{T}}$ whose projections on the bottom of $\mathcal{T}$ are contained in $[q^r, f^{h^l}(q^l)] \cup [f^{h^r}(q^r), q^l]$. Define also $\mathcal{F} = \mathcal{X} \cup \widetilde{\mathcal{D}}$ and $\mathcal{P} = \mathcal{F}^c$. See Figure 2. Denote by l_{ij}^l a generic l-interval that decomposes L_i, where the superscript refers to the side, left, of the interval, L_i. A missing l or r refers to either l or r. Define similarly l_{ij}^r, r_{ij}^l and r_{ij}^r. Since the towers are affine the l-intervals that enter into the decomposition of a fixed interval, say L_i, all have the same length which we denote by l_i^l (respectively for l_i^r, r_i^l and r_i^r). Since $\mathcal{T}$ is affine we have

$$\frac{n^l l_i^l}{L_i} = n^l \frac{l_i^l}{L_i} = n^l \frac{l_0^l}{L_0},$$

which shows that this quantity (respectively for $m^l r_i^l / L_i$, $n^r l_i^r / R_i$ and $m^r r_i^r / R_i$) depends only on the corresponding R_α-tower. Thus using Lemma 2, we can assume from now on that

$$\frac{1}{5} < \frac{m^l r_i^l}{n^l l_i^l} < \frac{m^r r_i^r}{n^r l_i^r} < 5. \tag{1}$$

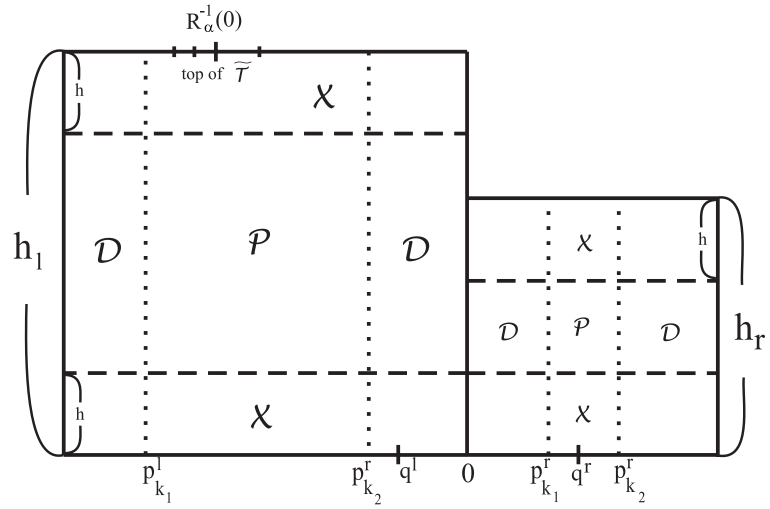

FIGURE 2.

Let $\{p^l_{ij}\}_{j=0}^{k^l}$, where $k^l = n^l + m^l$, be the extreme points of the partition of L_i by the l- and r-intervals. In a similar way define p^r_{ij} partitioning R_i. Now take k_1^r such that $p^r_{0k_1^r} < q^r \le p^r_{0k_1^r+1}$. Since $\|\widetilde{\mathcal{T}}\| \to 0$, as we cut and stack, on account of f being minimal, we see that $k_1^r \to \infty$ and $p^r_{0k_1^r} \to q^r$. In a similar way take positive integers k_2^r, k_1^l and k_2^l such that $p^l_{0k_1^l} < f^{h_r}(q^r) \le p^l_{0k_1^l+1}$, $p^l_{0k_2^l-1} < q^l \le p^l_{0k_2^l}$ and $p^r_{0k_2^r-1} < f^{h_l}(q^l) \le p^r_{0k_2^r}$, respectively.

Fix $0 < \xi < 1$. We are going to change $\widetilde{\mathcal{T}}$ to a pl tower, $\widetilde{\widetilde{\mathcal{T}}} = \widetilde{\widetilde{\mathcal{T}}}(h, \widetilde{h}, q^r, q^l, \xi)$, defining a pl homeomorphism $\widetilde{f}$ by contracting the intervals l^r_{ij} and r^l_{ij} inside the intervals of $\mathcal{T}$. We change the lengths of the l- and r-intervals by moving slightly the point p_{ij} to a point $\widetilde{p}_{ij}$. Start by changing the position of the points $f^i(p^r_{01})$, for $i = 0, 1, \ldots, h^r + h^l - 1$ (which are equal to p^r_{i1} or $p^l_{i-h^r 1}$), taking $\widetilde{p}^r_{01} = p^r_{01}$ and moving the point $f^i(p^r_{01})$, $i > 0$, to the point at distance $(p^r_{01} - p^r_{00})g^r_{00}g^r_{10} \cdots g^r_{i-10}$ from the nearest extreme (respectively p^r_{i0} or $p^l_{i-h^r 0}$) where $g^r_{s0} = \sqrt{D^+ f(p^r_{s0}) D^- f(p^r_{s0})}$ is the geometric mean of the lateral derivatives at the extreme points of the intervals. This changes the points next to the lower extremes of the L- and R-intervals. The points next to the upper extremes $f^i(p^l_{0k^l-1})$, for $i = 0, 1, \ldots, h^r + h^l - 1$, are similarly moved. Since we have kept the extremes of the large intervals fixed and changed the length of the small intervals next to them in order that both lateral derivatives are now equal to the geometric mean of the previous lateral derivatives, it follows that $\widetilde{\widetilde{f}}$ agrees with f and is smooth at the extremes of the intervals of $\mathcal{T}$. To move the remaining points p^r_{ij} and to define the one-floor-up maps in S, we use the previous lemma $h^r - 1$ times taking $I_1 = R_i$, $I_2 = R_{i+1}$, $k_1 = k_1^r$, $k_2 = k_2^r$, $\mathcal{Q} = \widetilde{\mathcal{X}} = \widetilde{\mathcal{X}}(\widetilde{h})$ and $\mathcal{C}$ the remaining l^r_{ij} intervals. The contraction λ_i^r will change as we move in the levels of the right tower as follows:

$$\lambda_i^r = \begin{cases} \xi^i, & \text{for } i = 0, 1, \ldots, h-1, \\ \xi^{h^r - i - 1}, & \text{for } i = h^r - h, h^r - h + 1, \ldots, h^r - 1, \\ \xi^h, & \text{otherwise.} \end{cases}$$

To move the points in the left tower of $\mathcal{T}$ we proceed in the same way, only now contracting the intervals r_{ij}^l not in $\widetilde{\mathcal{X}}$. Note that λ_{i+1}/λ_i is either 1 or $\xi^{\pm 1}$ and $0 < 1 - \xi \le \xi^{-1} - 1$.

Using (1) we see that we can take $\kappa = 6$ in Lemma 4. In fact, take $\widetilde{\mathcal{T}}$ further down in the sequence of pairs of towers, if necessary, in order to keep the ratios $\mu(\cup\mathcal{C})/\mu(\cup\mathcal{E})$ in $(1/6, 6) \supseteq (1/5, 5)$ which is possible since

$$\mu(\cup\widetilde{\mathcal{X}}) \le 4\widetilde{h}\|\widetilde{\mathcal{T}}\|$$

and $\|\widetilde{\mathcal{T}}\| \to 0$.

To define the top-to-bottom maps of $\widetilde{\widetilde{\mathcal{T}}}$ we use Lemma 4 again on the intervals $I_1 = L_{h^l-1}$, $I_2 = f(L_{h^l-1})$ (respectively $I_1 = R_{h^r-1}$, $I_2 = f(R_{h^r-1})$) with the k's as before and the partitions given by the points p_{ij} where we omit the points $0 = p_{00}^r = p_{0k^l}^l$, which has no pre-image, and $f^{\widetilde{a}q_{\widetilde{n}+1}+q_{\widetilde{n}}-1}(0)$, which has no image in the stretch of trajectory from 0 to $\widetilde{a}q_{\widetilde{n}+1} + q_{\widetilde{n}} - 1$ which we are considering. The sets $\mathcal{C}$ and $\mathcal{Q}$ will not, in fact, matter since we are going to take, as we must, $\lambda_1 = \lambda_2 = 1$ because we need $\widetilde{\widetilde{f}}$ to be an isometry near $R_\alpha^{-1}(0)$ in order to keep the rotation number the same. Recall that in Lemma 4 the points p_{ik_1} and p_{ik_2} are unchanged and from the definition of the q's and k_i's we have $R_\alpha^{-1}(0) \in f^{h^r-1}[p_{0k_1^r}^r, p_{0k_2^r}^r] \cup f^{h^l-1}[p_{0k_1^l}^l, p_{0k_2^l}^l]$.

The definition of $\widetilde{\widetilde{\mathcal{T}}}$ is now complete and we have the following lemma.

LEMMA 5. *Given $\mathcal{T}$, f, such that $1/2 < \|Df\| < 2$, with h, $\widetilde{h}$, q^l and q^r and ξ as above, there are towers $\widetilde{\mathcal{T}}$, infinitely often in the sequence of towers, which deformed to $\widetilde{\widetilde{\mathcal{T}}} = \widetilde{\widetilde{\mathcal{T}}}(h, \widetilde{h}, q^r, q^l, \xi)$ as described satisfy:*

(a) $\|D\widetilde{f} - Df\| \le \max\{\frac{2}{3}\|JDf\|, 4r\|Df\|, 6(\xi^{-1} - 1)\|Df\|\}$;

(b) $\|JDf\| \le \max\{\frac{2}{3}\|JDf\| + 4r\|Df\|, (6(\xi^{-1} - 1) + 4r)\|Df\|, 12(\xi^{-1} - 1)\|Df\|\}$;

(c) $\mu(\widetilde{S} - S) \le \mu(\cup\mathcal{F}) + \xi^h\mu(S) + \mu(\cup\widetilde{\mathcal{X}})$;

(d) $\mu(\widetilde{S}) \ge (1 - \xi^h)\mu(S) - \mu(\cup\mathcal{F}) - \mu(\cup\widetilde{\mathcal{X}})$;

(e) $\rho(\widetilde{\widetilde{f}}) = \alpha$ *and* $\widetilde{\widetilde{f}}^i(0) = f^i(0), i = 0, \ldots, aq_{n+1} + q_n - 1$;

where

$$r = \max\left\{\frac{\widetilde{p}_{i1} - p_{i0}}{p_{ik_1} - p_{i0}}, \frac{p_{ik} - \widetilde{p}_{ik-1}}{p_{ik} - p_{ik_2}}\right\},$$

with the maximum taken over all intervals in $\mathcal{T}$. (a) and (b) hold for $D\widetilde{f}^{-1}$.

Proof. (e) is obvious from the choice of the points q and from the definition of the new top to bottom maps. (a) and (b) follow easily from Lemma 4 and its following remark plus the easy fact that the geometric mean, $\sqrt{\rho\lambda}$, of two numbers ρ and λ in the interval $(1/2, 2)$

satisfy $\sqrt{\rho\lambda}-\rho<2/3(\lambda-\rho)$ and $\lambda-\sqrt{\rho\lambda}<2/3(\lambda-\rho)$. (c) follows from

$$\mu(\widetilde{S}-S)=\sum_{i,j}\widetilde{r}_{ij}^{l}\le\mu(\cup\mathcal{F})+\sum_{\widetilde{r}_{ij}^{l}\in\mathcal{P}}\widetilde{r}_{ij}^{l}\le\mu(\cup\mathcal{F})+\xi^{h}\mu(S)+\mu(\cup\widetilde{\mathcal{X}})$$

and (d) from

$$\begin{aligned}\mu(\widetilde{S})&=\sum_{i,j}\widetilde{r}_{ij}^{l}+\sum_{i,j}\widetilde{r}_{ij}^{r}\ge\sum_{i,j}\widetilde{r}_{ij}^{r}=\mu(S)-\sum_{i,j}\widetilde{l}_{ij}^{r}\\&\ge\mu(S)-\mu(\cup\mathcal{F})-\sum_{\widetilde{l}_{ij}^{r}\in\mathcal{P}}\widetilde{l}_{ij}^{r}\ge\mu(S)-\mu(\cup\mathcal{F})-\xi^{h}\mu(S)-\mu(\cup\widetilde{\mathcal{X}}).\end{aligned}$$ □

4. *Proof of Theorem 1*

We use the notation of Lemma 5. Fix ν, $0<\nu<1/2$, and make $\sigma_0=0$ and $\sigma_n=\sum_{i=1}^{n}\nu^i$. Take a and b, $0<a<1/6$ and $5/6<b<1$, and fix a positive integer n_0 such that $a+\nu^{n_0}<1/6$, $b+\nu^{n_0}<1$ and $2\nu^{n_0}<\nu$. Take a sequence x_n, $n=0,1,2,\ldots$, such that $0<x_{n+1}<x_n$ and $3/4\,x_n<x_{n+1}$, $\forall n$ and $\sum x_n=1$. Define positive integers

$$h_n=1+\left[\frac{\ln(\nu^{n+n_0}(\nu-2\nu^{n_0}))}{\ln(100+x_{n+1}/2)-\ln(100+x_{n+1})}\right],\quad n=0,1,2,\ldots.$$

Then there is a sequence f_n, $n=0,1,2,\ldots$, of pl homeomorphisms defined by a sequence of pairs of towers $\mathcal{T}_n$ such that, denoting by S_n the union of the intervals in the right tower of $\mathcal{T}_n$, we have:

(1) $\rho(f_n)=\alpha$, $\forall n\ge 0$;
(2) $f_n=f_{n-1}$ on the extremes of $\mathcal{T}_{n-1}$, $\forall n\ge1$;
(3) $\|Df_n-Df_{n-1}\|$ and $\|Df_n^{-1}-Df_{n-1}^{-1}\|\le x_n$, $\forall n\ge1$;
(4) $\|JDf_n\|$ and $\|JDf_n^{-1}\|\le x_n$, $\forall n\ge1$;
(5) $a+\nu^{n+n_0}\le\mu(S_n)\le b+\nu^{n_0}\sigma_n$, $\forall n\ge0$;
(6) $\mu(S_n-S_{n-1})\le\nu^{n+n_0}$, $\forall n\ge1$;
(7) h_n^l and $h_n^r>2h_n$, $\forall n\ge0$, where h_n^l and h_n^r denote the heights of the left and right towers of $\mathcal{T}_n$, respectively;
(8) $\mu(\cup\mathcal{X}(h_n))\le\frac{1}{2}\nu^{n+2n_0}$, $\forall n\ge0$;
(9) $\|\mathcal{T}_n\|\le x_n$, $\forall n\ge0$.

We construct the sequences f_n and $\mathcal{T}_n$ by induction on $n=0,1,\ldots$.

For $n=0$ we take a tower $\mathcal{T}_0$ of R_α. More precisely, according to Lemma 2, constructing the towers of R_α, we obtain $1/5<(1-\mu(S_0))/\mu(S_0)<5$ or $1/6<\mu(S_0)<5/6$ infinitely often, where S_0 is the union of the intervals in the right tower of $\mathcal{T}_0$. Fix one of these towers that is high and thin enough to satisfy (7), (8) and (9) as $\mathcal{T}_0$.

Now suppose we have constructed towers $\mathcal{T}_0,\mathcal{T}_1,\ldots,\mathcal{T}_n$ defining f_0, $f_1,\ldots,f_n$, respectively, pl homeomorphisms satisfying (1) to (9). (3) implies

$$\|Df_n\|\quad\text{and}\quad\|Df_n^{-1}\|\le1+\sum_{i=1}^{n}x_i<2.$$

Define $\mathcal{T}_{n+1}=\widetilde{\widetilde{\mathcal{T}}}$ given by the previous lemma where we take $\mathcal{T}=\mathcal{T}_n$, $h=h_n$, $\widetilde{h}=h_{n+1}$, $\xi=(100+x_{n+1}/2)/(100+x_{n+1})$ (then $\xi^{-1}-1<x_{n+1}/100$ and $\xi^{h_n}<\nu^{n+n_0}(\nu-2\nu^{n_0})$), q^l and q^r so close to zero that

(i)

$$\mu(\cup\mathcal{D}(q^l, q^r)) < \frac{1}{2}\nu^{n+1+2n_0} \quad \text{and} \quad |q| < \frac{x_{n+1}}{2^{h_n^l+h_n^r}},$$

and $\widetilde{\mathcal{T}}$ so high in the sequence of towers that

(ii)

$$\|\widetilde{\mathcal{T}}\| < \min\left\{x_{n+1}, \frac{\nu^{n+1+2n_0}}{8h_{n+1}}\right\},$$

(iii) $\widetilde{h}^l$ and $\widetilde{h}^r > 2h_{n+1}$,

(iv)

$$\max\left\{\frac{p_{i1}-p_{i0}}{p_{ik_1}-p_{i0}}, \frac{p_{ik}-p_{ik-1}}{p_{ik}-p_{ik_2}}\right\} < \frac{x_{n+1}}{100}2^{-2(h_n^r+h_n^l)},$$

for both right and left p's. This is possible since, for instance, $p_{i1}^l \to p_{i0}^l$ and $p_{ik_1^l}^l \to f^{h_r+i}(q^r)$ as we move up in the sequence of towers:

(v) $\mu(\cup\widetilde{\mathcal{X}}(h_{n+1})) \le \frac{1}{2}\nu^{n+1+2n_0}$.

Using (iv) and the definition of the points $\widetilde{p}_{ij}$ next to the extremes we see that

$$r = \max\left\{\frac{\widetilde{p}_{i1}-p_{i0}}{p_{ik_1}-p_{i0}}, \frac{p_{ik}-\widetilde{p}_{ik-1}}{p_{ik}-p_{ik_2}}\right\} < \frac{x_{n+1}}{100}$$

since, for instance,

$$\frac{\widetilde{p}_{i1}^r - p_{i0}^r}{p_{ik_1}^r - p_{i0}^r} = \frac{(p_{01}^r - p_{00}^r)g_{00}^r g_{10}^r \cdots g_{i-10}^r}{p_{ik_1}^r - p_{i0}^r}$$

$$= \frac{p_{01}^r - p_{00}^r}{p_{0k_1}^r - p_{00}^r}\frac{\sqrt{D^- f^i(p_{00}^r)}}{\sqrt{D^+ f^i(p_{00}^r)}} \le \frac{x_{n+1}}{100}2^{-2(h_n^r+h_n^l)}2^{2i} < \frac{x_{n+1}}{100}.$$

We check that the assertions (1)–(9) hold. (1) and (2) follow from (e). (3) is obtained from Lemma 5(a) since

$$\tfrac{2}{3}\|JDf_n\| \le \tfrac{2}{3}x_n < x_{n+1}$$
$$4r\|Df_n\| \le 8r \le x_{n+1}$$
$$6(\xi^{-1}-1)\|Df_n\| \le 12(\xi^{-1}-1) < x_{n+1}.$$

(4) follows from Lemma 5(b) and the definitions. The same estimates hold for Df_{n+1}^{-1} and JDf_{n+1}^{-1}.

Checking (6) we have

$$\begin{aligned}\mu(S_{n+1}-S_n) &\le \mu(\cup\mathcal{F}_n) + \xi^{h_n}\mu(S_n) + \mu(\cup\mathcal{X}_{n+1})\\ &\le \mu(\cup\mathcal{D}) + \mu(\cup\mathcal{X}_n) + \xi^{h_n} + \mu(\cup\mathcal{X}_{n+1})\\ &\le \tfrac{1}{2}\nu^{n+1+2n_0} + \tfrac{1}{2}\nu^{n+2n_0} + \nu^{n+2n_0}(\nu - 2\nu^{2n_0}) + \tfrac{1}{2}\nu^{n+1+2n_0} < \nu^{n+1+n_0}.\end{aligned}$$

The second inequality of (5) is easy to prove:

$$\mu(S_{n+1}) \le \mu(S_n) + \mu(S_{n+1}-S_n) \le b + \nu^{n_0}\sigma_n + \nu^{n+1+n_0} = b + \nu^{n_0}\sigma_{n+1}.$$

The first is a bit longer:

$$\mu(S_{n+1}) \geq \mu(S_n) - \mu(\cup\mathcal{F}_n) - \xi^{h_n}\mu(S_n) - \mu(\cup\mathcal{X}_{n+1}) \geq a + \nu^{n+n_0}$$
$$- (\tfrac{1}{2}\nu^{n+1+2n_0} + \tfrac{1}{2}\nu^{n+2n_0} + \nu^{n+2n_0}(\nu - 2\nu^{2n_0}) + \tfrac{1}{2}\nu^{n+1+2n_0})$$
$$\geq a + \nu^{n+1+n_0}.$$

Condition (7) is (iii) and (8) is (v).

(9) is clearly true for the l- or r-intervals outside the intervals $[\widetilde{p}_{ik_1}, \widetilde{p}_{ik-k_2}]$ since, for example,

$$\widetilde{p}^{r}_{ik_1} - \widetilde{p}^{r}_{i0} < f_n^i(q^r) - p^r_{i0} = Df_n^i(p_{00})(q^r - p_{00}) < 2^i|q^r|$$

which is less than x_{n+1} by (i), and also for the intervals there which are contracted or remain with the same length on account of (ii). If the interval is expanded, say an interval r^r_{ij} in R_i, denoting this expansion by ν_i we have $M\nu_i r^r_i < R_i \leq \|\mathcal{T}_n\| \leq x_n$, where M is the number of intervals expanded, but then $\nu_i r^r_i \leq x_n/M^r < x_n/2 < x_{n+1}$, completing the induction.

Now (2) implies that h_n converges pointwise on $\mathcal{O}$ to an order-preserving map h and, similarly, f_n converges on $h(\mathcal{O})$, a set dense in $\mathbb{T}^1$, to an order-preserving map f. Using Herman's Lemma 3.3 [**2**, p. 140], these maps extend to homeomorphisms of $\mathbb{T}^1$, also denoted by h and f, respectively, such that $h \circ R_\alpha = f \circ h$. That f is C^1 follows from Lemma 3 and (3) and (4). It remains to see that f is not ergodic. Denote by $\mathcal{T}^0_n$ the pair of towers of R_α corresponding to $\mathcal{T}_n$ and, similarly, write $S^0_n = h_n^{-1}(S_n) = h^{-1}(S_n)$ as its right tower. The set $\mathcal{S} = \bigcap_{n=0}^\infty \bigcup_{k=n}^\infty S^0_k$ is R_α invariant and therefore $h(\mathcal{S}) = \bigcap_{n=0}^\infty \bigcup_{k=n}^\infty S_k$ is f invariant. But

$$\mu(h(\mathcal{S})) = \limsup_{n\to\infty}(\mu(S_n) + \mu(S_{n+1} - S_n) + \cdots)$$
$$\leq \limsup_{n\to\infty}(\mu(S_n) + \nu^{n+1+n_0} + \cdots)$$
$$\leq \limsup_{n\to\infty}\mu(S_n) \leq \limsup_{n\to\infty} b + \nu^{n_0}\sigma_n < b + \nu^{n_0} < 1.$$

On the other hand,

$$\mu(h(\mathcal{S})) = \lim_{n\to\infty}\mu\bigg(\bigcup_{k=n}^\infty S_k\bigg) \geq \liminf_{n\to\infty}\mu(S_n) \geq \liminf_{n\to\infty} a + \nu^{n+n_0} = a > 0. \qquad \square$$

References

[1] A. Denjoy. Sur les courbes définies par les équations différentielles à la surface du tore. *J. Math. Pures Appl.* **9**(11) (1932), 333–375.

[2] M. R. Herman. Sur la conjugaison différentiable des difféomorphismes du circle à des rotations. *Publ. Math. IHES* **49** (1979), 5–234.

[3] Y. Katznelson and D. Ornstein. The differentiability of the conjugation of certain diffeomorphisms of the circle. *Ergod. Th. & Dynam. Sys.* **9** (1989), 643–680.

[4] W. de Melo and S. van Strien. *One Dimensional Dynamics*. Springer, New York, 1993.