Commun. Math. Phys. 112, 89–101 (1987)

Communications in
Mathematical
Physics

A New Proof of M. Herman's Theorem

K. M. Khanin and Ya. G. Sinai

L. D. Landau Institute for Theoretical Physics, Kosygin Str. 2, SU-117334 Moscow, V-334, USSR

Dedicated to Walter Thirring on his 60th birthday

Abstract. A new proof of the M. Herman theorem on the smooth conjugacy of a circle map is presented here. It is based on the thermodynamic representation of dynamical systems and the study of the ergodic properties for the corresponding radom variables.

1. Introduction

In this paper we present a new proof of M. Herman's famous theorem about smooth conjugacy of diffeomorphisms of the circle to rotations. Our proof is based on a version of thermodynamic formalism which was used earlier for the study of Feigenbaum's mappings of the interval (see [1]) and later for some critical mappings of the circle (see [2, 3]).

Consider a strictly monotone continuous function φ such that $\varphi(x+1) = \varphi(x)+1$. It defines a homeomorphism of S^1 through the equation: $T_\varphi x = \{\varphi(x)\}$, $x \in [0;1)$. Denote the rotation number of T_φ by ϱ.

Assumptions. A.1. $\varphi \in C^{2+\gamma}$, $\gamma > 0$, $\varphi' \geqq \text{const} > 0$;

A.2. ϱ is irrational and if

$$\varrho = [k_1, k_2, \dots, k_n, \dots]$$

is the expansion of ϱ into continued fraction, then $k_n \leqq \text{const}\, n^\nu$, $\nu > 0$.

If ϱ is irrational and $\varphi \in C^2$, then Denjoy's classical theorem states that T_φ is topologically isomorphic to the rotation with angle ϱ. In other words there exists a strictly monotone continuous function ψ, $\psi(x+1) = \psi(x)+1$, such that $\psi(\varphi(x)) = \psi(x)+\varrho$. If we denote $R_\varrho = T_{\varphi_0}$ $\varphi_0 = x+\varrho$, then the last equality means $T_\psi T_\varphi = R_\varrho T_\psi$.

Herman's Theorem. *Under the Assumptions A.1, A.2 the function* $\psi \in C^1$.

For us it will be convenient to prove a statement equivalent to Herman's theorem. This is pointed out in Arnold's paper [11]. Let us write the equation for the density $\pi(x)$ of an invariant absolutely continuous measure, provided that such a measure exists:

$$\frac{\pi(T_\varphi(x))}{\pi(x)} = \frac{1}{\varphi'(x)}. \tag{1}$$

Herman's Theorem. *Under the Assumptions A.1, A.2, Eq.* (1) *has a continuous strictly positive solution.*

The equivalence of the two formulations follows from the equality $\psi(x) = \int_0^x \pi(y)dy$. Take any point $x_0 \in S^1$ and its semi-trajectory $O = \{x_n\}_0^\infty$, $x_n = T_\varphi^n x_0$ which is everywhere dense on S^1. Put $\pi(x_0) = 1$ and $\pi(x_i) = \prod_{k=0}^{i-1} (\varphi'(x_k))^{-1}$. Thus, the function π is defined on O and satisfies there (1). We shall show that it is continuous on O and can be continued to a positive continuous function on the whole circle.

Our method uses the ideas of the renormalization group theory in the theory of dynamical systems, as it is presented in [4, 5]. We hope that it can also be useful for the critical mappings of the circle.

The first proof of Herman's theorem was published in [6]. Later there appeared shorter versions (see [7]). For us the exposition presented in the dissertation by Yoccoz [8] was very useful.

In contrast to the positive results of this paper, there are examples [12] of C^2 diffeomorphisms arbitrarily close to a rotation which are not C^1 conjugate to a rotation, even if their rotation numbers satisfy the hypothesis given here.

2. The Symbolic Representations of Real Numbers Generated by T_φ

Denote by $\varrho_n = \frac{p_n}{q_n}$ the n^{th} approximant of ϱ, i.e. $\varrho_n = [k_1, k_2, \ldots, k_n]$. The numbers p_n, q_n satisfy the recurrence relations

$$p_{n+1} = k_{n+1}p_n + p_{n-1}, \qquad p_1 = 1, \qquad p_0 = 0, \tag{2'}$$

$$q_{n+1} = k_{n+1}q_n + q_{n-1}, \qquad q_1 = k_1, \qquad q_0 = 1. \tag{2''}$$

Take any integer $m \leqq q_{n+1}$. It can be written in the form

$$m = a_n q_n + a_{n-1}q_{n-1} + \ldots + a_0 q_0,$$

where $0 \leqq a_i \leqq k_{i+1}$. This representation is nonunique. For any $x_0 \in S^1$, let us denote by $C_0^{(n)}(x_0)$ the closed intervals whose end-points are x_0, x_{q_n}, $C_i^{(n)}(x_0) = T_\varphi^i C_0^{(n)}(x_0)$. The following lemma is well-known.

Lemma 1. *For every n the intervals* $C_i^{(n)}(x_0)$, $0 \leqq i < q_{n+1}$, $C_j^{(n+1)}(x_0)$, $0 \leqq j < q_n$ *are mutually disjoint except the end-points, and cover the whole circle.*

Introduce the partition ξ_n of S^1 whose elements are $C_i^{(n)}(x_0)$, $0 \leqq i \leqq q_{n+1} - 1$, $C_j^{(n+1)}(x_0)$, $0 \leqq j \leqq q_n - 1$. The partitions ξ_n are monotone-increasing, $\xi_n \leqq \xi_{n+1}$ in the sense that each C_{ξ_n} consists of several $C_{\xi_{n+1}}$. To be more precise each $C_j^{(n+1)}(x_0)$, $0 \leqq j < q_n$ is also an element of the partition ξ_{n+1}. But each $C_i^{(n)}(x_0)$ is partitioned into $(k_{n+2}+1)$ elements of ξ_{n+1} because we have the equality:

$$C_i^{(n)}(x_0) = C_i^{(n+2)}(x_0) \cup \bigcup_{j=0}^{k_{n+2}-1} C_{i+q_n+jq_{n+1}}^{(n+1)}(x_0). \tag{3}$$

In other words if $O_n = \bigcup_{i=0}^{q_{n+1}+q_n-1} x_i$, then each $C_i^{(n)}$ contains exactly k_{n+2} new points of O_{n+1}. It will also be important that $C_i^{(n)}(x_0) = C_0^{(n)}(x_i)$. One can construct a unique symbolic representation of any $x \in S^1$ not belonging to the set of end-points of all $C_i^{(n)}(x_0)$, $0 \leqq i < q_{n+1}$, using some inductive procedure. Put $a_{n+1} = A$ if $x \in C_i^{(n+1)}(x_0)$, $0 \leqq i < q_n$ and $a_{n+1} = k_{n+2} - j$ if $x \in C_{i+q_n+jq_{n+1}}^{(n+1)}(x_0) \subset C_i^{(n)}(x_0)$, $0 \leqq i < q_{n+1}$, $0 \leqq j < k_{n+2}$ [see (3)]. For

$$x \in C_i^{(n)}(x_0) - \bigcup_{j=0}^{k_{n+2}-1} C_{i+q_n+jq_{n+2}}^{(n+1)}(x_0) = C_i^{(n+2)}(x_0) \quad \text{put} \quad a_{n+1} = 0 .$$

Thus, we write

$$x \leftrightarrow (a_0, a_1, \ldots, a_n, \ldots) = a ,$$

where $a_i = A,\ 0, 1, \ldots, k_{i+1},\ i \neq 0,\ a_0 = A,\ 0, \ldots, k_1 - 1$. The Lebesgue measure ℓ induces a probability measure on the space of all admissible sequences $a = (a_0, a_1, \ldots, a_n, \ldots)$. It is easy to see that the only restriction on the word a is the following one: $a_{n+1} = A$ iff $a_n = 0$. This is a "hard core" type condition in statistical mechanics. Remark that each element of the partition ξ_n corresponds under this representation to a finite word $(a_0, a_1, \ldots, a_n)$ of length $(n+1)$. In particular, the words $(0, A, \ldots, 0, A)$ and $(A, 0, \ldots, A, 0)$ correspond to the elements $C_0^{(n)}$, $C_0^{(n+1)}$, respectively. The proof of Herman's theorem will be based on the ergodic properties of the sequence of random variables a_n with respect to the measure induced by the Lebesgue measure.

3. The Formulation of the Main Lemmas and the Derivation of Herman's Theorem

Let us recall the

Denjoy Lemma (see [9]). *There exists an absolute constant $C > 0$ such that for every $x_0 \in S^1$, $n > 0$,*

$$e^{-C} \leqq \prod_{i=0}^{q_n-1} \varphi'(x_i) \leqq e^C .$$

Henceforth all constants depend on φ but do not depend on n. Assume that Denjoy's lemma can be strengthened in the following way:

$$e^{-\varepsilon_n} \leqq \prod_{i=0}^{q_n-1} \varphi'(x_i) \leqq e^{\varepsilon_n} , \tag{4}$$

where $\sum_n k_{n+1}\varepsilon_n < \infty$. Then it implies Herman's theorem. Indeed, let us consider $\pi(x_i)$ for $0 \leqq i < q_{n+1}$. Take $C_i^{(n)}$, $0 \leqq i < q_{n+1}$, whose end-points are x_i, x_{i+q_n}. It contains inside itself k_{n+2} points $x_{i+q_n+jq_{n+1}}$, $1 \leqq j \leqq k_{n+2}$, and

$$e^{-\varepsilon_{n+1}} \leqq \frac{\pi(x_{i+q_n+jq_{n+1}})}{\pi(x_{i+q_n+(j-1)q_{n+1}})} = \left(\prod_{t=i+q_n+(j-1)q_{n+1}}^{i+q_n+jq_{n+1}-1} \varphi'(x_t) \right)^{-1} \leqq e^{\varepsilon_{n+1}}$$

$$e^{-\varepsilon_n} \leqq \frac{\pi(x_{i+q_n})}{\pi(x_i)} \leqq e^{\varepsilon_n} .$$

It shows that $\pi(x_{i+q_n+jq_{n+1}})$ differs from $\pi(x_i)$ by a multiplicative factor which is bounded by $e^{\pm(k_{n+2}\varepsilon_{n+1}+\varepsilon_n)}$. Thus, for every $x_j \in C_i^{(n)}$, we have

$$\exp\left\{-\varepsilon_n - \sum_{m \geqq n} (k_{m+2}+1)\varepsilon_{m+1}\right\} \leqq \frac{\pi(x_j)}{\pi(x_i)}$$

$$\leqq \exp\left\{\varepsilon_n + \sum_{m \geqq n} (k_{m+2}+1)\varepsilon_{m+1}\right\},$$

which implies the desired continuity and positivity of π on O. Now we formulate some lemmas which will give a stronger version of Denjoy's lemma.

For any $y \in S^1$ consider $y \in C_{i_0}^{(n)}(x_0) \subset C_{i_k}^{(n-k)}(x_0)$.

Lemma 2. *There exists a constant $\lambda < 1$ not depending on y, k, n, x_0 such that*

$$\ell(C_{i_0}^{(n)}(x_0)) \leqq \lambda^k \ell(C_{i_k}^{(n-k)}(x_0)).$$

Put $H_m(y_0) = \sum_{i=0}^{q_m-1} \log \varphi'(y_i)$, and take $y_0^{(1)}, y_0^{(2)} \in C_i^{(n)}(x_0)$.

Lemma 3. $|H_{n-k}(y_0^{(1)}) - H_{n-k}(y_0^{(2)})| \leqq \text{const}\,\lambda^k$, *where the constant λ is the same as in Lemma 2.*

Put $\Delta^{(n)}x_j = |x_{j+q_{n+1}} - x_j|$, and

$$b_n = \max\left\{\max_{0 \leqq p < q_n, x_0} \left|\sum_{j=0}^{p} \frac{\varphi''(x_j)}{\varphi'(x_j)} \Delta^{(n)}x_j\right|;\ \lambda^{n\gamma}\right\}.$$

Lemma 4. $|H_n(y_0) - H_n(x_0)| \leqq \text{const} \cdot b_n$ *for every* $y_0 \in C_0^{(n+1)}(x_0)$.

Lemma 5. $|H_n(y_0) - H_n(x_0)| \leqq \text{const} \sum_{k=0}^{[\frac{n}{2}]} b_{n-k}(n-k)^\nu + \text{const}\, n^{\nu+1}\lambda^{n/2}$ *for arbitrary* $x_0, y_0 \in S^1$.

Lemma 6.

$$\left|\sum_{j=0}^{q_n-1} \frac{\varphi''(x_j)}{\varphi'(x_j)} \Delta^{(n)}x_j\right| \leqq \text{const}\,\lambda_1^{\sqrt{n}}, \qquad \lambda < \lambda_1 < 1.$$

Lemma 7. *For every* $0 \leqq p < q_n$,

$$\left|\sum_{j=0}^{p} \frac{\varphi''(x_j)}{\varphi'(x_j)} \Delta^{(n)}x_j\right| \leqq \text{const}\, n^{2\nu+2}\lambda_2^{\sqrt{n}}, \qquad \lambda_1 < \lambda_2 < 1.$$

Lemmas 4–7 give the stronger version of Denjoy's lemma alluded to before. We may choose $\varepsilon_n = \text{const}\, n^{3\nu+3}\lambda_2^{\sqrt{n}}$ in (4). All lemmas are proved in the next two sections.

4. Proof of Lemmas 2–5

Proof of Lemma 2. Take an element $C_i^{(n-1)}$ of the partition $\xi^{(n-1)}$. The partition $\xi^{(n)}$ decomposes $C_i^{(n-1)}$ into $(k_{n+1}+1)$ elements. One of them is $C_i^{(n+1)}$, while the others belong to the trajectory of the $C_0^{(n)}$. Take first $C_i^{(n+1)}$. There exists some

$C_j^{(n)} \subset C_i^{(n-1)}$ such that $C_{j+q_n}^{(n)} \supset C_i^{(n+1)}$. By Denjoy's lemma

$$e^{-C} \leqq \frac{\ell(C_{j+q_n}^{(n)})}{\ell(C_j^{(n)})} \leqq e^C$$

and

$$\frac{\ell(C_i^{(n+1)})}{\ell(C_i^{(n-1)})} \leqq \frac{\ell(C_i^{(n+1)})}{\ell(C_j^{(n)})+\ell(C_i^{(n+1)})} \leqq \frac{1}{1+\dfrac{\ell(C_j^{(n)})}{\ell(C_{j+q_n}^{(n)})}} \leqq (1+e^{-C})^{-1}.$$

Assume now that $k_{n+1}>1$ and take some $C_j^{(n)} \subset C_i^{(n-1)}$. Then $C_{j\pm q_n}^{(n)} \subset C_i^{(n-1)}$ for at least one choice of sign. Then again with the help of Denjoy's lemma

$$\frac{\ell(C_j^{(n)})}{\ell(C_i^{(n-1)})} \leqq \frac{\ell(C_j^{(n)})}{\ell(C_j^{(n)})+\ell(C_{j\pm q_n}^{(n)})} \leqq (1+e^{-C})^{-1}.$$

Now we discuss the case $k_{n+1}=1$. Here $C_i^{(n-1)} = C_i^{(n+1)} + C_{i+q_{n-1}}^{(n)}$ and the ratio $\dfrac{\ell(C_{i+q_{n-1}}^{(n)})}{\ell(C_i^{(n-1)})}$ can be very close to 1. The partition $\xi^{(n+1)}$ decomposes $C_{i+q_{n-1}}^{(n)}$ into $(k_{n+2}+1)$ elements. As above the ratio of the lengths of these elements to $\ell(C_i^{(n-1)})$ is less than $(1+e^{-C})^{-1}$ if $k_{n+2}>1$. If $k_{n+2}=1$, then we have $C_{i+q_{n-1}+q_n}^{(n+1)} \subset C_{i+q_{n-1}}^{(n)} \subset C_i^{(n-1)}$ and

$$\frac{\ell(C_{i+q_{n-1}+q_n}^{(n+1)})}{\ell(C_i^{(n-1)})} \leqq \frac{\ell(C_{i+q_{n-1}+q_n}^{(n+1)})}{\ell(C_i^{(n+1)})+\ell(C_{i+q_{n-1}+q_n}^{(n+1)})} \leqq (1+e^{-2C})^{-1}.$$

Thus, after two steps the ratio in all cases becomes less than $(1+e^{-2C})^{-1}$. This immediately gives the statement of the lemma.

Proof of Lemma 3. Take $y_0^{(1)}, y_0^{(2)} \subset C_i^{(n)}(x_0) \subset C_j^{(n-k)}(x_0)$ or $C_j^{(n-k+1)}(x_0)$. First we remark that $C_i^{(n)}(x_0) = C_0^{(n)}(x_i)$. Thus, we may replace x_0 by x_i and assume that $y_0^{(1)}, y_0^{(2)} \in C_0^{(n)}(x) \subset C_0^{(n-m)}(x)$, where $m=k$ or $k-1$ is an even number. Then $y_i^{(1)}, y_i^{(2)} \in C_i^{(n)}(x) \subset C_i^{(n-m)}(x)$. Now we can write using Lemma 2,

$$|H_{n-k}(y_0^{(1)}) - H_{n-k}(y_0^{(2)})| \leqq \sum_{i=0}^{q_{n-k}-1} |\log\varphi'(y_i^{(1)}) - \log\varphi'(y_i^{(2)})|$$

$$\leqq \max_{x\in S^1} \left|\frac{\varphi''(x)}{\varphi'(x)}\right| \cdot \max_i \frac{\ell(C_i^{(n)}(x))}{\ell(C_i^{(n-m)}(x))} \leqq \mathrm{const}\,\lambda^m. \quad \text{QED.}$$

Proof of Lemma 4. Let $I = \sum_{i=0}^{q_n-1} [\log\varphi'(y_i) - \log\varphi'(x_i)]$. We have

$$I = \sum_{i=0}^{q_n-1} \frac{\varphi''(x_i)}{\varphi'(x_i)}(y_i - x_i) + \sum_{i=0}^{q_n-1} \alpha_i (y_i - x_i)^{1+\gamma} = I_1 + I_2, \ |\alpha_i| \leqq \mathrm{const}.$$

For I_2 we have a trivial estimate with the help of Lemma 2 for $k=n$:

$$|I_2| \leqq \mathrm{const}\max|y_i - x_i|^\gamma \leqq \mathrm{const}\,\lambda^{n\gamma}.$$

Introduce $\varrho_i = \dfrac{y_i - x_i}{\Delta^{(n)} x_i}$.

$$\varrho_{i+1} = \frac{\varphi'(x_i)(y_i - x_i) + \frac{1}{2}\varphi''(x_i)(y_i - x_i)^2 + \alpha_i^{\mathrm{I}}(y_i - x_i)^{2+\gamma}}{\varphi'(x_i)\Delta^{(n)}x_i + \frac{1}{2}\varphi''(x_i)(\Delta^{(n)}x_i)^2 + \alpha_i^{\mathrm{II}}(\Delta^{(n)}x_i)^{2+\gamma}}$$

$$= \varrho_i \frac{1 + \frac{1}{2}\dfrac{\varphi''(x_i)}{\varphi'(x_i)}(y_i - x_i) + \alpha_i^{\mathrm{III}}(y_i - x_i)^{1+\gamma}}{1 + \frac{1}{2}\dfrac{\varphi''(x_i)}{\varphi'(x_i)}\Delta^{(n)}x_i + \alpha_i^{(\mathrm{IV})}(\Delta^{(n)}x_i)^{1+\gamma}}.$$

In the above (and henceforth) $\alpha_i^{\mathrm{I}}, \alpha_i^{\mathrm{II}}, \ldots$ are some uniformly bounded numbers. Put

$$\varrho_i' = \varrho_i \prod_{j<i}\left[1 + \frac{1}{2}\frac{\varphi''(x_j)}{\varphi'(x_j)}\Delta^{(n)}x_j + \alpha_i^{(\mathrm{IV})}(\Delta^{(n)}x_j)^{1+\gamma}\right] = \varrho_i(1 + b_i').$$

For b_i' we have the estimate $|b_i'| \leqq \mathrm{const}\, b_n$. Furthermore,

$$\varrho_{i+1}' = \varrho_i'\left[1 + \frac{1}{2}\frac{\varphi''(x_i)}{\varphi'(x_i)}\varrho_i'(1 + b_i')^{-1}\Delta^{(n)}x_i + \alpha_i^{\mathrm{III}}(\varrho_i'(1 + b_i')^{-1})^{1+\gamma}(\Delta^{(n)}x_i)^{1+\gamma}\right]$$

and

$$\frac{1}{\varrho_{i+1}'} = \frac{1}{\varrho_i'}\left[1 - \frac{1}{2}\frac{\varphi''(x_i)}{\varphi'(x_i)}\varrho_i'(1 + b_i')^{-1}\Delta^{(n)}x_i + \alpha_i^{(\mathrm{V})}(\varrho_i'(1 + b_i')^{-1})^{1+\gamma}(\Delta^{(n)}x_i)^{1+\gamma}\right],$$

$$\frac{1}{\varrho_{i+1}'} = \frac{1}{\varrho_i'} - \frac{1}{2}\frac{\varphi''(x_i)}{\varphi'(x_i)}(1 + b_i')^{-1}\Delta^{(n)}x_i + \alpha_i^{(\mathrm{V})}(\varrho_i')^{\gamma}(1 + b_i')^{-1-\gamma}(\Delta^{(n)}x_i)^{1+\gamma}.$$

This yields

$$\left|\frac{1}{\varrho_i'} - \frac{1}{\varrho_0'}\right| \leqq \mathrm{const}\left|\sum_{j<i}\frac{\varphi''(x_j)}{\varphi'(x_j)}\Delta^{(n)}x_j\right| + \mathrm{const}\, b_n \leqq \mathrm{const}\, b_n,$$

and

$$|\varrho_0' - \varrho_i'| \leqq \varrho_0'\varrho_i' \,\mathrm{const}\, b_n \leqq \mathrm{const}\, b_n \varrho_0,$$

$$|\varrho_0 - \varrho_i| \leqq \mathrm{const}\, b_n \varrho_0.$$

The last inequality shows that ϱ_i differs from ϱ_0 by a very small number (assuming that b_n are small enough), i.e. the mapping $T_\varphi^i : C_0^{(n+1)} \to C_i^{(n+1)}$ can be well enough approximated by a linear map. Now we can write

$$|I_1| = \left|\sum_{i=0}^{q_n-1}\frac{\varphi''(x_i)}{\varphi'(x_i)}(y_i - x_i)\right| = \left|\sum_{i=0}^{q_n-1}\frac{\varphi''(x_i)}{\varphi'(x_i)}\Delta^{(n)}x_i\varrho_i\right|$$

$$\leqq \varrho_0\left|\sum_{i=0}^{q_n-1}\frac{\varphi''(x_i)}{\varphi'(x_i)}\Delta^{(n)}x_i\right| + \left|\sum_{i=0}^{q_n-1}\frac{\varphi''(x_i)}{\varphi'(x_i)}\Delta^{(n)}x_i(\varrho_i - \varrho_0)\right|$$

$$\leqq \mathrm{const}\, b_n\varrho_0. \quad \text{QED.}$$

Proof of Lemma 5. Take x_0 and construct the sequence of the intervals $C_i^{(m)}(x_0)$, $m=n, n+1$. Then $y_0 \in C_i^{(n+1)}(x_0)$ for some i, $0 \leqq i < q_n$, or $y_0 \in C_j^{(n)}(x_0)$ for some j, $0 \leqq j < q_{n+1}$. In the first case

$$J_n = H_n(y_0) - H_n(x_0) = \sum_{s=0}^{q_n-1-i} [\log\varphi'(y_s) - \log\varphi'(x_{s+i})] + \sum_{s=q_n-i}^{q_n-1} [\log\varphi'(y_s) - \log\varphi'(x_{s-q_n+i})].$$

Remark that $y_{q_n-i} \in C_{q_n}^{(n+1)}(x_0) \subset C_0^{(n)}(x_0)$ and recall (see Sect. 2) that $C_j^{(m)}(x_i) = C_{i+j}^{(m)}(x_0)$. Therefore, each of the sums has the form $J'_r = \sum_{s=0}^{r} [\log\varphi'(z_s^{(1)}) - \log\varphi'(z_s^{(2)})]$, where $z_0^{(1)}, z_0^{(2)} \in C_0^{(n)}(x)$ for some x and $0 \leqq r \leqq q_n - 1$. Write (see Sect. 2)

$$r = a_{n-1}q_{n-1} + a_{n-2}q_{n-2} + \ldots + a_0 q_0,$$

where $0 \leqq a_i \leqq k_{i+1}$. Then

$$J'_r = \sum_{j:a_j \neq 0} \sum_{i=1}^{a_j} \sum_{\substack{a_{n-1}q_{n-1}+\ldots+a_{j+1}q_{j+1}+(i-1)q_j \leqq s \\ < a_{n-1}q_{n-1}+\ldots+a_{j+1}q_{j+1}+iq_j}} [\log\varphi'(z_s^{(1)}) - \log\varphi'(z_s^{(2)})].$$

Consider each inner sum separately. For every s the points $z_s^{(1)}, z_s^{(2)}$ belong to some $C_0^{(n)}$. Using Lemma 4 for long cycles $\left(j \geqq \frac{n}{2}\right)$ and Lemma 3 for short cycles $\left(j < \frac{n}{2}\right)$ we can write:

$$|J'_r| = \left|\sum_{j \geqq \frac{n}{2}}\right| + \left|\sum_{j < \frac{n}{2}}\right| \leqq \text{const} \sum_{k=0}^{\left[\frac{n}{2}\right]} b_{n-k}(n-k)^{\nu} + \text{const}\, n^{\nu+1}\lambda^{n/2},$$

which implies the desired result.

In the second case the arguments are the same if $j < q_n$. If $q_n \ell \leqq j < q_n(\ell+1)$, $0 < \ell \leqq k_{n+1}$, we apply the same arguments for the differences,

$$\sum_{s=0}^{q_n-1} [\log\varphi'(y_{s-tq_n}) - \log\varphi'(y_{s-(t-1)q_n})], \qquad 1 \leqq t \leqq \ell,$$

and then consider the difference

$$\sum_{s=0}^{q_n-1} [\log\varphi'(x_s) - \log\varphi'(y_{s-\ell q_n})].$$

Our previous arguments also work now, and thus, we get the needed estimation. QED.

Corollary. $|H_n(x_0)| = \left|\sum_{i=0}^{q_n-1} \log\varphi'(x_i)\right| \leqq \text{const}\, n^{3\nu+3}\lambda_2^{\sqrt{n}}$.

Indeed, apply Lemmas 5–7 for x_0 and y_0, where $H_n(y_0) = 0$. Such y_0 exists, because

$$\underbrace{\varphi(\varphi(\ldots\varphi}_{q_n \text{ times}}(x+1)\ldots) - \underbrace{\varphi(\varphi(\ldots\varphi}_{q_n \text{ times}}(x)\ldots) = 1,$$

and in view of the mean value theorem there exists y_0 for which

$$\frac{d}{dx}\underbrace{\varphi(\varphi\ldots\varphi(y_0)\ldots)}_{q_n\text{ times}}=\prod_{i=0}^{q_n-1}\varphi'(y_i)=1, \qquad H_n(y_0)=0.$$

We have used the results of Lemmas 6 and 7. Their proofs do not depend on Lemmas 4 and 5.

5. Proofs of Lemmas 6 and 7

Proof of Lemma 6. We have to estimate the sum

$$I_n=\sum_{j=0}^{q_n-1}\frac{\varphi''(x_j)}{\varphi'(x_j)}\Delta^{(n)}x_j.$$

Putting $k=\left[\frac{n}{2}\right]$, rewrite it in the following way:

$$\begin{aligned}I_n=&\sum_{s=0}^{q_{n-k+1}-1}\ \sum_{j:C_j^{(n+1)}\subset C_s^{(n-k)}}\frac{\varphi''(x_j)}{\varphi'(x_j)}\Delta^{(n)}x_j\\&+\sum_{t=0}^{q_{n-k}-1}\ \sum_{j:C_j^{(n+1)}\subset C_t^{(n-k+1)}}\frac{\varphi''(x_j)}{\varphi'(x_j)}\Delta^{(n)}x_j.\end{aligned}$$

Choose an arbitrary point $y_s^{(n-k)}$, $y_t^{(n-k+1)}$ in each $C_s^{(n-k)}, C_t^{(n-k+1)}$. Then

$$\begin{aligned}I_n=&\sum_{s=0}^{q_{n-k+1}-1}\frac{\varphi''(y_s^{(n-k)})}{\varphi'(y_s^{(n-k)})}\ell(C_s^{(n-k)})p_s^{(n-k)}\\&+\sum_{t=0}^{q_{n-k}-1}\frac{\varphi''(y_t^{(n-k+1)})}{\varphi'(y_t^{(n-k+1)})}\ell(C_t^{(n-k+1)})p_t^{(n-k+1)}\\&+\sum_{s=0}^{q_{n-k+1}-1}\sum_j\left(\frac{\varphi''(x_j)}{\varphi'(x_j)}-\frac{\varphi''(y_s^{(n-k)})}{\varphi'(y_s^{(n-k)})}\right)\Delta^{(n)}x_j\\&+\sum_{t=0}^{q_{n-k}-1}\sum_j\left(\frac{\varphi''(x_j)}{\varphi'(x_j)}-\frac{\varphi''(y_t^{(n-k+1)})}{\varphi'(y_t^{(n-k+1)})}\right)\Delta^{(n)}x_j,\end{aligned}\tag{5}$$

where

$$p_s^{(n-k)}=\frac{1}{\ell(C_s^{(n-k)})}\sum_{j:C_j^{(n+1)}\subset C_s^{(n-k)}}\ell(C_j^{(n+1)}),$$

$$p_t^{(n-k+1)}=\frac{1}{\ell(C_t^{(n-k+1)})}\sum_{j:C_j^{(n+1)}\subset C_t^{(n-k+1)}}\ell(C_j^{(n+1)}).$$

Due to the smooth properties of φ each of the two last sums is not more than $\mathrm{const}\,\lambda^{\frac{n\gamma}{2}}$.

Assume that we have succeeded in proving that for some constant p the differences

$$|p_s^{(n-k)}-p|,\ |p_t^{(n-k+1)}-p|\leqq\mathrm{const}\,\lambda_1^{\sqrt{n}}.\tag{6}$$

Then

$$\sum_{s=0}^{q_{n-k+1}-1} \frac{\varphi''(y_s^{(n-k)})}{\varphi'(y_s^{(n-k)})} \ell(C_s^{(n-k)}) p_s^{(n-k)} + \sum_{t=0}^{q_{n-k}-1} \frac{\varphi''(y_t^{(n-k+1)})}{\varphi'(y_t^{(n-k+1)})}$$

$$\cdot p_t^{(n-k+1)} \ell(C_t^{(n-k+1)}) = p \left[\sum_{s=0}^{q_{n-k+1}-1} \frac{\varphi''(y_s^{(n-k)})}{\varphi'(y_s^{(n-k)})} \ell(C_s^{(n-k)}) \right.$$

$$\left. + \sum_{t=0}^{q_{n-k}-1} \frac{\varphi''(y_t^{(n-k+1)})}{\varphi'(y_t^{(n-k+1)})} \ell(C_t^{(n-k+1)}) \right] + \sum_{s=0}^{q_{n-k+1}-1} \frac{\varphi''(y_s^{(n-k)})}{\varphi'(y_s^{(n-k)})}$$

$$\cdot \ell(C_s^{(n-k)}) (p_s^{(n-k)} - p) + \sum_{t=0}^{q_{n-k}-1} \frac{\varphi''(y_t^{(n-k+1)})}{\varphi'(y_t^{(n-k+1)})} \ell(C_t^{(n-k+1)}) (p_t^{(n-k+1)} - p).$$

The absolute value of the last two sums is not more than const $\lambda_1^{\sqrt{n}}$. The expression in square brackets is an approximation of the Riemann integral $\int \frac{\varphi''(y)}{\varphi'(y)} dy = 0$. Due to the smoothness properties of the function φ it differs from the integral by a number with absolute value less than const $\lambda^{\frac{n\gamma}{2}}$. Thus, we have to show only inequalities (6). It will be proved via some ergodic theorem type arguments for the Markov chains.

Let us return to the increasing sequence of partitions ξ_n and the corresponding symbolic representation which were introduced in Sect. 2. Consider the conditional probabilities $\ell(a_n | a_{n-1}, \dots, a_0) = \frac{\ell(a_0, a_1, \dots, a_n)}{\ell(a_0, a_1, \dots, a_{n-1})}$, wehere $\ell(a_0, \dots, a_m)$ means the Lebesgue measure of the element of ξ_m corresponding to the admissible word $(a_0, \dots, a_m)$.

Lemma 8. *The following inequality holds:*

$$e^{-\text{const}\,\lambda^s} \leqq \frac{\ell(a_n | a_{n-1}, \dots, a_{n-s}, a''_{n-s-1}, \dots, a''_0)}{\ell(a_n | a_{n-1}, \dots, a_{n-s}, a'_{n-s-1}, \dots, a'_0)} \leqq e^{\text{const}\,\lambda^s},$$

provided both words are admissible.

Proof. The words $(a'_0, \dots, a'_{n-s-1}, a_{n-s}, \dots, a_n)$, $(a'_0, \dots, a'_{n-s-1}, a_{n-s}, \dots, a_{n-1})$, $(a'_0, \dots, a'_{n-s-1})$ correspond to the intervals $C_{i_0}^{(m_0)} \subset C_{i_1}^{(m_1)} \subset C_{i_2}^{(m_2)}$, where $m_0 = n, n+1$, $m_1 = n-1, n$, $m_2 = n-s-1, n-s$. Then the words $(a''_0, \dots, a''_{n-s-1}, a_{n-s}, \dots, a_n)$, $(a''_0, \dots, a''_{n-s-1}, a_{n-s}, \dots, a_{n-1})$, $(a''_0, \dots, a''_{n-s-1})$ correspond to the intervals $C_{i_0+j}^{(m_0)} \subset C_{i_1+j}^{(m_1)} \subset C_{i_2+j}^{(m_2)}$, where $i_2 + j < q_{n-s-1}$, if $m_2 = n-s$ and $i_2 + j < q_{n-s}$ if $m_2 = n-s-1$. Denote

$$\varrho_k = \frac{\ell(C_{i_0+k}^{(m_0)})}{\ell(C_{i_1+k}^{(m_1)})} : \frac{\ell(C_{i_0}^{(m_0)})}{\ell(C_{i_1}^{(m_1)})}, \qquad 1 \leqq k \leqq j.$$

We have

$$\varrho_{k+1} = \frac{\int_{C_{i_0+k}^{(m_0)}} \varphi'(y)\,dy}{\int_{C_{i_1+k}^{(m_1)}} \varphi'(y)\,dy} : \frac{\ell(C_{i_0}^{(m_0)})}{\ell(C_{i_1}^{(m_1)})}.$$

Using the mean value theorem we can write

$$\varrho_{k+1}=\frac{\varphi'(y_{i_0+k}^{(m_0)})}{\varphi'(y_{i_1+k}^{(m_1)})}\varrho_k\,,$$

where $y_i^{(m)}$ are some points of the corresponding intervals. Furthermore,

$$\exp\{-\mathrm{const}\,\ell(C_{i_1+k}^{(m_1)})\}\leqq\frac{\varphi'(y_{i_0+k}^{(m_0)})}{\varphi'(y_{i_1+k}^{(m_1)})}\leqq\exp\{\mathrm{const}\,\ell(C_{i_1+k}^{(m_1)})\}\,.$$

Then

$$\frac{\ell(C_{i_0+j}^{(m_0)})}{\ell(C_{i_1+j}^{(m_1)})}=\varrho_j\frac{\ell(C_{i_0}^{(m_0)})}{\ell(C_{i_1}^{(m_1)})}$$

and

$$\exp\left\{-\mathrm{const}\sum_{t=0}^{j-1}\ell(C_{i_1+t}^{(m_1)})\right\}\leqq\varrho_j\leqq\exp\left\{\mathrm{const}\sum_{t=0}^{j-1}\ell(C_{i_1+t}^{(m_1)})\right\}.$$

However, due to Lemma 2,

$$\sum_{t=0}^{j-1}\ell(C_{i_1+t}^{(m_1)})\leqq\sum_{t=0}^{j-1}\ell(C_{i_2+t}^{(m_2)})\frac{\ell(C_{i_1+t}^{(m_1)})}{\ell(C_{i_2+t}^{(m_2)})}\leqq\mathrm{const}\,\lambda^s.\quad \text{QED.}$$

Lemma 8 shows that the sequence of random variables $\{a_n\}$ can be well enough approximated by the Markov chain. Consider now the conditional probabilities $\ell(a_{n+m},\dots,a_n|a'_{n-1},\dots,a'_0)$, $\ell(a_{n+m},\dots,a_n|a''_{n-1},\dots,a''_0)$, provided both words

$$(a'_0,\dots,a'_{n-1},a_n,\dots,a_{n+m}),\ (a''_0,\dots,a''_{n-1},a_n,\dots,a_{n+m})$$

are admissible.

Lemma 9. *There exists an absolute constant $C_1>0$ such that for all n,m,*

$$e^{-C_1}\leqq\frac{\ell(a_{n+m},\dots,a_n|a'_{n-1},\dots,a'_0)}{\ell(a_{n+m},\dots,a_n|a''_{n-1},\dots,a''_0)}\leqq e^{C_1}\,.$$

The proof follows easily from Lemma 8 and the equations

$$\ell(a_{n+m},\dots,a_n|a'_{n-1},\dots,a'_0)=\prod_{i=0}^{m}\ell(a_{n+i}|a_{n+i-1},\dots,a'_{n-1},\dots,a'_0)\,,$$

$$\ell(a_{n+m},\dots,a_n|a''_{n-1},\dots,a''_0)=\prod_{i=0}^{m}\ell(a_{n+i}|a_{n+i-1},\dots,a''_{n-1},\dots,a''_0)\,.$$

The proof of the following lemma is also simple.

Lemma 10. *There exists an absolute constant $C_2>0$ such that for all n,m and all the admissible words $(a'_0,\dots,a'_{n-3})$, $(a''_0,\dots,a''_{n-3})$, $(a_n,\dots,a_{n+m})$:*

$$e^{-C_2}\leqq\frac{\ell(a_{n+m},\dots,a_n|a'_{n-3},\dots,a'_0)}{\ell(a_{n+m},\dots,a_n|a''_{n-3},\dots,a''_0)}\leqq e^{C_2}\,.$$

Return now to $p_s^{(n-k)}, p_t^{(n-k+1)}$. In terms of symbolic representation they can be written in the form:

$$p_s^{(n-k)},p_t^{(n-k+1)}=\ell(a_n=0|a_{n-k},\dots,a_0)\,.$$

We have to estimate the difference

$$|\ell(a_n=0|a_{n-k},\ldots,a_0)-\ell(a_n=0)|.$$

Lemma 11. $|\ell(a_n=0|a_{n-k},\ldots,a_0)-\ell(a_n=0)|\leqq \text{const}\,\lambda_3^{\sqrt{k}}$ *for some constant* $\lambda_3<1$.

Lemma 11 is a Markov ergodic theorem and it can be proved by the methods of the Markov chain theory. This technique is well-known (see [10]); thus, we shall describe only the main steps.

Fix an integer m, $m\sim\sqrt{k}$ and introduce a new probability measure on the words

$$\tilde{a}=(a_n,a_{n-1},\ldots,a_{n-m+3},a_{n-m},\ldots,a_{n-2m+3},$$
$$a_{n-2m},\ldots,a_{n-3m+3},\ldots,a_{n-(i-1)m},\ldots,a_{n-im+3},a_{n-im},\ldots,a_0)$$

by the formula

$$\ell'(\tilde{a})=\ell(a_0,\ldots,a_{n-im})\ell(a_{n-(i-1)m},\ldots,a_{n-im+3}|a_{n-im},\ldots,a_0)$$
$$\cdot\prod_{j=0}^{i-2}\ell(a_{n-jm},\ldots,a_{n-(j+1)m+3}|a_{n-(j+1)m},\ldots,a_{n-(j+2)m+3}).$$

Here $i\sim\sqrt{k}$. This measure is the one of the Markov chain with memory m. It follows easily from Lemma 8 that

$$\exp(-\text{const}\,\lambda^m\cdot i)\leqq\frac{\ell'(\tilde{a})}{\ell(\tilde{a})}\leqq\exp(\text{const}\,\lambda^m\cdot i).$$

If we consider the Markov transition operator corresponding to ℓ' for the transition to m steps, then it follows from Lemma 10 that it is a contraction with a coefficient uniformly less than 1. Then the usual ergodic theorem for Markov chains shows that the difference between the conditional probabilities $\ell(a_n|a_{n-im},\ldots,a_0)$ for different $a_{n-im},\ldots,a_0$ is less than λ_4^i, $\lambda_4<1$. This gives the desired result. QED.

Proof of Lemma 7. Lemma 7 is derived from Lemma 6 in a similar way as Lemma 5 is derived from Lemma 4. Consider the sums

$$I_{n-2i}=\sum_{j=0}^{q_{n-2i}-1}\frac{\varphi''(x_j)}{\varphi'(x_j)}\Delta^{(n)}x_j.$$

As in the proof of Lemma 6 consider

$$p_s^{(n-k)}=\frac{1}{\ell(C_s^{(n-k)})}\sum_{\substack{j:C_j^{(n+1)}\subset C_s^{(n-k)}\\ 0\leqq j<q_{n-2i}}}\ell(C_j^{(n+1)}),$$

$$p_t^{(n-k+1)}=\frac{1}{\ell(C_t^{(n-k+1)})}\sum_{\substack{j:C_j^{(n+1)}\subset C_t^{(n-k+1)}\\ 0\leqq j<q_{n-2i}}}\ell(C_j^{(n+1)}).$$

It is easy to see that in the symbolic representation $p_s^{(n-k)}, p_t^{(n-k+1)}$ has the following form:

$$p_s^{(n-k)},p_t^{(n-k+1)}=\ell(\underbrace{0,A,0,A,\ldots,0}_{2i+1}|a_{n-k},\ldots,a_0).$$

Then the same arguments as in the proof of Lemma 6 show that

$$|I_{n-2i}| \leqq \text{const}\, \lambda_1^{\sqrt{n}} \quad \text{for} \quad i < \frac{m}{2} \sim \sqrt{\frac{n}{8}}.$$

Moreover, for short cycles we have the trivial estimate

$$|I_{n-j}| \leqq \text{const}\, \lambda^j, j \geqq m \sim \sqrt{\frac{n}{2}}.$$

Take now $p \leqq q_n$. We can decompose p in the following way,

$$p = \sum_{i=1}^{\left[\frac{m}{2}\right]} a_i q_{n-2i} + \sum_{j=m}^{n} b_j q_{n-j},$$

where $a_i \leqq \text{const}\, n^{2\nu+1}$, $b_m \leqq \text{const}\, n^{2\nu+1}$, $b_j \leqq \text{const}(n-j)^\nu$, $j > m$. Using the above estimates we obtain

$$\left| \sum_{j=0}^{p} \frac{\varphi''(x_j)}{\varphi'(x_j)} \Delta^{(n)} x_j \right| \leqq \left(\sum_{i=1}^{\left[\frac{m}{2}\right]} a_i \right) \text{const}\, \lambda_1^{\sqrt{n}} + \text{const} \sum_{j=m}^{n} b_j \lambda^j.$$

This estimate gives the desired result. QED.

Remarks. 1. Using a more sophisticated technique one can show that in Lemma 11 the right-hand side can be taken as const $\cdot\, \lambda_3^k$. This makes it possible to prove an exponential estimate in Denjoy's lemma and also to consider the rotation numbers ϱ, $\varrho = [k_1, \ldots, k_n, \ldots]$, for which $k_n \leqq \text{const} \cdot B^n$ with some $B > 1$ as in [8]. It also implies that conjugacy $\psi \in C^{1+\tau}$, for some $\tau > 0$.

2. From the viewpoint of the renormalization group ideology the result proved here means the convergence to the linear fixed point of the renormalization group transformation (see [4]) for nondegenerate diffeomorphism of the circle.

References

1. Vul, E.B., Sinai, Ya.G., Khanin, K.M.: Feigenbaum universality and the thermodynamic formalism. Usp. Mat. Nauk. **39**:3, 3–37 (1984) [English transl. Russ. Math. Surv. **39**, 1–40 (1984)]
2. Halsey, T.C., Jensen, M.H., Kadanoff, L.P., Procaccia, I., Shraiman, B.I.: Fractal measures and their singularities: the characterization of strange sets. Phys. Rev. A **33**, 2, 1141–1151 (1986)
3. Collet, P., Lebowitz, J., Porzio, A.: Dimension spectrum for some dynamical systems. Rutgers Univ. Preprint
4. Khanin, K.M., Sinai, Ya.G.: Renormalization group method and the K.A.M. theory. In: Nonlinear phenomena in plasma physics and hydrodynamics. Sagdeev, R.Z. (ed.). Moscow: Mir 1986
5. Sinai, Ya.G., Khanin, K.M.: Renormalization group method in the theory of dynamical systems. In: Proceedings of the conference "Renormalization Group 86", Dubna 1986 (in press)
6. Herman, M.: Sur la conjugaison différentiable des difféomorphism du cercle à des rotations. Pub. Mat. I.H.E.S. **49**, 5–233 (1979)
7. Herman, M.: Simple proofs of local conjugacy theorems for diffeomorphisms of the circle with almost every rotation number. Bol. Soc. Bras. Mat. **16**, 45–83 (1985);

Sue les difféomorphisms du cercle de nombre de rotation de type constant. In: Papers in Honor of A. Zygmund. Becker, A. et al. (eds.). Belmont: Wadsworth 1983

8. Yoccoz, J.C.: Centralisateurs et conjugaison différentiable des diffeomorphisms du cercle. Thesis Univ. Paris Sud (1985) unpublished;
C^1 conjugaison des difféomorphisms du cercle. In: Geometry and dynamic. Palis, J. (ed.). Lecture Notes in Mathematics, Vol. 1007. Berlin, Heidelberg, New York: Springer 1983;
Conjugaison différentiable des difféomorphisms du cercle dont le nombre de rotation vérifie une condition diophantienne. Ann. Sci. Éc. Norm. Super. **17**, 333–359 (1984)
9. Cornfeld, I.P., Sinai, Ya.G., Fomin, S.V.: Ergodic theory. Moscow: Nauka 1980 (English transl.: Berlin, Heidelberg, New York: Springer 1982)
10. Ruelle, D.: Thermodynamic Formalism. Reading: Addison-Wesley 1978
11. Arnol'd, V.I.: Small denominators. I. Mappings of the circumference onto itself. Izv. Mat. Nauk. **25**:1, 25–96 (1961) [English Transl. Am. Math. Soc. **49**, 213–284]
12. Hawkins, J., Schmidt, K.: On C^2 diffeomorphisms of the circle which are of type III_1. Invent. math. **66**, 511–518 (1966)

Communicated by E. Lieb

Received January 15, 1987