Herman’s Theory Revisited (Extension)

A. Teplinsky∗

February 1, 2008

Abstract

We prove that a $C^{3+\beta}$-smooth orientation-preserving circle diffeomorphism with rotation number in Diophantine class $D_\delta$, $0 < \beta < \delta < 1$, is $C^{2+\beta-\delta}$-smoothly conjugate to a rigid rotation.

1 Introduction

An irrational number $\rho$ is said to belong to Diophantine class $D_\delta$ if there exists a constant $C > 0$ such that $|\rho - p/q| \geq Cq^{-2-\delta}$ for any rational number $p/q$.

In [1], the following result was proven.

Theorem (Khanin-T.). Let $T$ be a $C^{2+\alpha}$-smooth orientation-preserving circle diffeomorphism with rotation number $\rho \in D_\delta$, $0 < \delta < \alpha \leq 1$. Then $T$ is $C^{1+\alpha-\delta}$-smoothly conjugate to the rigid rotation by angle $\rho$.

By the smoothness of conjugacy we mean the smoothness of the homeomorphism $\phi$ such that

$$\phi \circ T \circ \phi^{-1} = R_\rho,$$

(1)

where $R_\rho(\xi) = \xi + \rho \mod 1$ is the mentioned rigid rotation.

The aim of the present paper is to extend the Theorem above to the case of $T \in C^{3+\beta}$, $0 < \beta < \delta < 1$, so that the extended result is read as follows:

Theorem 1. Let $T$ be a $C^r$-smooth orientation-preserving circle diffeomorphism with rotation number $\rho \in D_\delta$, $0 < \delta < 1$, $2 + \delta < r < 3 + \delta$. Then $T$ is $C^{r-1-\delta}$-smoothly conjugate to the rigid rotation by angle $\rho$.

Historically, the first global results on smoothness of conjugation with rotations were obtained by M. Herman [2]. Later J.-C. Yoccoz extended the theory to the case of Diophantine rotation numbers [3]. The result, recognized generally as the final answer in the theory, was proven by Y. Katznelson, D. Ornstein [4]. In our terms it states that the conjugacy is $C^{r-1-\delta-\varepsilon}$-smooth for any $\varepsilon > 0$ provided that $0 < \delta < r - 2$. Notice that Theorem 1 is stronger than the result just cited, though valid for a special scope of parameter values only, and it is sharp, i.e. smoothness of conjugacy higher than $C^{r-1-\delta}$ cannot be achieved in general settings, as it

∗Institute of Mathematics, Kiev, Ukraine
follows from the examples constructed in [4]. At present, we do not know whether Theorem 1 can be extended further, and the examples mentioned do not prevent such an extension.

In paper by K. Khanin, Ya. Sinai [5], published simultaneously with [4], similar problems were approached by a different method. The method we use is different from the one of [4]; it is based on the ideas of [5], the cross-ratio distortion tools and certain exact relations between elements of the dynamically generated structure on the circle.

All the implicit constants in asymptotics written as $O(\cdot)$ depend on the function $f$ only in Section 2 and on the diffeomorphism $T$ only in Section 3.

## 2 Cross-ratio tools

The **cross-ratio** of four pairwise distinct points $x_1, x_2, x_3, x_4$ is

$$\text{Cr}(x_1, x_2, x_3, x_4) = \frac{(x_1 - x_2)(x_3 - x_4)}{(x_2 - x_3)(x_4 - x_1)}$$

Their **cross-ratio distortion** with respect to a strictly increasing function $f$ is

$$\text{Dist}(x_1, x_2, x_3, x_4; f) = \frac{\text{Cr}(f(x_1), f(x_2), f(x_3), f(x_4))}{\text{Cr}(x_1, x_2, x_3, x_4)}$$

Clearly,

$$\text{Dist}(x_1, x_2, x_3, x_4; f) = \frac{D(x_1, x_2, x_3; f)}{D(x_1, x_4, x_3; f)}, \quad (2)$$

where

$$D(x_1, x_2, x_3; f) = \frac{f(x_1) - f(x_2)}{x_1 - x_2} : \frac{f(x_2) - f(x_3)}{x_2 - x_3}$$

is the **ratio distortion** of three distinct points $x_1, x_2, x_3$ with respect to $f$.

In the case of smooth $f$ such that $f'$ does not vanish, both the ratio distortion and the cross-ratio distortion are defined for points, which are not necessarily pairwise distinct, as the appropriate limits (or, just by formally replacing ratios $(f(a) - f(a))/(a - a)$ with $f'(a)$ in the definitions above).

Notice that both ratio and cross-ratio distortions are multiplicative with respect to composition: for two functions $f$ and $g$ we have

$$D(x_1, x_2, x_3; f \circ g) = D(x_1, x_2, x_3; g) \cdot D(g(x_1), g(x_2), g(x_3); f) \quad (3)$$

$$\text{Dist}(x_1, x_2, x_3, x_4; f \circ g) = \text{Dist}(x_1, x_2, x_3, x_4; g) \cdot \text{Dist}(g(x_1), g(x_2), g(x_3), g(x_4); f) \quad (4)$$

For $f \in C^{3+\beta}$ it is possible to evaluate the next entry in the asymptotical expansions for both ratio and cross-ratio distortions. The **Swartz derivative** of $C^{3+\beta}$-smooth function is defined as $Sf = \frac{f'''}{f''} - \frac{3}{2} (\frac{f'''}{f''})^2$.

**Proposition 1.** Let $f \in C^{3+\beta}$, $\beta \in [0, 1]$, and $f' > 0$ on $[A, B]$. Then for any $x_1, x_2, x_3 \in [A, B]$ the following estimate holds:

$$D(x_1, x_2, x_3; f) = 1 + (x_1 - x_3) \left( \frac{f''(x_1)}{2f'(x_1)} + \frac{1}{6} Sf(x_1)(x_2 + x_3 - 2x_1) + O(\Delta^{1+\beta}) \right), \quad (5)$$

where $\Delta = \max\{x_1, x_2, x_3\} - \min\{x_1, x_2, x_3\}$. 

2
We start by proving the following

**Lemma 1.** For arbitrary \( \theta \in [A, B] \) we have

\[
\frac{f''(\theta)}{2f'(\theta)} + \frac{f'''(\theta)}{6f'(\theta)}(x_1 + x_2 + x_3 - 3\theta) - \left( \frac{f''(\theta)}{2f'(\theta)} \right)^2 (x_2 + x_3 - 2\theta) = \frac{f''(x_1)}{2f'(x_1)} + \frac{1}{6} S f(x_1)(x_2 + x_3 - 2x_1) + O(\Delta_\theta^{1+\beta}), \tag{6}
\]

where \( \Delta_\theta = \max\{x_1, x_2, x_3, \theta\} - \min\{x_1, x_2, x_3, \theta\} \).

**Proof.** Obvious estimates \( f''(x_1) = f''(\theta) + f''(\theta)(x_1 - \theta) + O(|x_1 - \theta|^{1+\beta}) \) and \( f'(x_1) = f'(\theta) + f''(\theta)(x_1 - \theta) + O((x_1 - \theta)^2) \) imply that

\[
\frac{f''(x_1)}{2f'(x_1)} = \frac{f''(\theta)}{2f'(\theta)} + \left( \frac{f''(\theta)}{2f'(\theta)} \right)^2 (x_1 - \theta) + O(\Delta_\theta^{1+\beta}). \tag{7}
\]

On the other hand, \( S f(x_1) = S f(\theta) + O(|x_1 - \theta|^\beta) \) and \( |x_2 + x_3 - 2x_1| \leq 2\Delta_\theta \), hence

\[
\frac{1}{6} S f(x_1)(x_2 + x_3 - 2x_1) = \left( \frac{f''(\theta)}{6f'(\theta)} - \frac{(f''(\theta))^2}{4(f'(\theta))^2} \right) (x_2 + x_3 - 2x_1) + O(\Delta_\theta^{1+\beta}) \tag{8}
\]

Adding (7) and (8) gives (6). \( \square \)

**Remark 1.** Notice, that Lemma 1 in particular, provides an alternative, more general (though less memorizable) formulation of Proposition 1 as we may choose \( \theta = x_2 \), or \( x_3 \), or any other point between \( \min\{x_1, x_2, x_3\} \) and \( \max\{x_1, x_2, x_3\} \) to get the same order \( O(\Delta^{1+\beta}) \) as in (3).

**Proof of Proposition 1** Using \( x_2 \) as the reference point for taking derivatives, we get

\[
\frac{f(x_1) - f(x_2)}{x_1 - x_2} = f'(x_2) + \frac{1}{2} f''(x_2)(x_1 - x_2) + \frac{1}{6} f'''(x_2)(x_1 - x_2)^2 + O(|x_1 - x_2|^{2+\beta}),
\]

\[
\frac{f(x_2) - f(x_3)}{x_2 - x_3} = f'(x_2) + \frac{1}{2} f''(x_2)(x_3 - x_2) + \frac{1}{6} f'''(x_2)(x_3 - x_2)^2 + O(|x_3 - x_2|^{2+\beta}),
\]

and after dividing (in view of the expansion \( (1 + t)^{-1} = 1 - t + t^2 + O(t^3) \)) obtain

\[
D(x_1, x_2, x_3; f) = 1 + (x_1 - x_3) \left[ \frac{f''(x_2)}{2f'(x_2)} + \frac{f'''(x_2)}{6f'(x_2)} (x_1 + x_3 - 2x_2) - \left( \frac{f''(x_2)}{2f'(x_2)} \right)^2 (x_3 - x_2) \right] + O(\Delta^{2+\beta}) \tag{9}
\]

In the case when \( x_2 \) lies between \( x_1 \) and \( x_3 \), the estimate (4) implies

\[
D(x_1, x_2, x_3; f) = 1 + (x_1 - x_3) \left[ \frac{f''(x_2)}{2f'(x_2)} + \frac{f'''(x_2)}{6f'(x_2)} (x_1 + x_3 - 2x_2) - \left( \frac{f''(x_2)}{2f'(x_2)} \right)^2 (x_3 - x_2) + O(\Delta^{1+\beta}) \right] \tag{10}
\]
It is not hard to notice that the expression in the square brackets here is exactly the subject of Lemma 1 with \( \theta = x_2 \), thus (5) is proven.

Suppose that \( x_1 \) lies between \( x_2 \) and \( x_3 \). Then the version of (5) for \( D(x_1, x_2, x_3; f) \) is proven. Also, the version of (9) for \( D(x_1, x_3, x_2; f) \) is proven. One can check the following exact relation takes place:

\[
D(x_1, x_2, x_3; f) = 1 + \frac{x_1 - x_3}{x_2 - x_3}(D(x_2, x_1, x_3; f) - 1)D(x_1, x_3, x_2; f)
\]

(11)

Substituting

\[
D(x_2, x_1, x_3; f) - 1 = (x_2 - x_3)
\left( \frac{f''(x_2)}{2f'(x_2)} + \frac{1}{6}sf(x_2)(x_1 + x_3 - 2x_2) + O(\Delta^{1+\beta}) \right)
\]

and

\[
D(x_1, x_3, x_2; f) = 1 + (x_1 - x_2)\frac{f''(x_2)}{2f'(x_2)} + O(\Delta^{1+\beta})
\]

into (11), we get (10), and Lemma 1 again implies (5).

The case when \( x_3 \) lies between \( x_1 \) and \( x_2 \) is similar to the previous one. The case when two or three among the points \( x_1, x_2 \) and \( x_3 \) coincide, all considerations above are valid with obvious alterations.

**Proposition 2.** Let \( f \in C^{3+\beta}, \beta \in [0, 1] \), and \( f' > 0 \) on \([A, B]\). For any \( x_1, x_2, x_3, x_4 \in [A, B] \) the following estimate holds:

\[
\text{Dist}(x_1, x_2, x_3, x_4; f) = 1 + (x_1 - x_3)(\frac{1}{6}(x_2 - x_3)Sf(\theta) + O(\Delta^{1+\beta}))
\]

(12)

where \( \Delta = \max\{x_1, x_2, x_3, x_4\} - \min\{x_1, x_2, x_3, x_4\} \) and \( \theta \) is an arbitrary point between \( \min\{x_1, x_2, x_3, x_4\} \) and \( \max\{x_1, x_2, x_3, x_4\} \).

**Proof.** Proposition 1 and Lemma 1 imply

\[
D(x_1, x_2, x_3; f) = 1 + (x_1 - x_3)\left[ \frac{f''(\theta)}{2f'(\theta)} + \frac{f''(\theta)}{6f'(\theta)}(x_1 + x_2 + x_3 - 3\theta) \right. \\
\left. - \left( \frac{f''(\theta)}{2f'(\theta)} \right)^2 (x_2 + x_3 - 2\theta) + O(\Delta^{1+\beta}) \right]
\]

\[
D(x_1, x_4, x_3; f) = 1 + (x_1 - x_3)\left[ \frac{f''(\theta)}{2f'(\theta)} + \frac{f''(\theta)}{6f'(\theta)}(x_1 + x_4 + x_3 - 3\theta) \right. \\
\left. - \left( \frac{f''(\theta)}{2f'(\theta)} \right)^2 (x_4 + x_3 - 2\theta) + O(\Delta^{1+\beta}) \right]
\]

Dividing the first expression by the second one accordingly to (2) in view of the formula \((1 + t)^{-1} = 1 - t + t^2 + O(t^3)\), we get (12).

**Remark 2.** Obviously enough, the estimate (12) can be re-written as

\[
\log \text{Dist}(x_1, x_2, x_3, x_4; f) = (x_1 - x_3)\left( \frac{1}{6}(x_2 - x_3)Sf(\theta) + O(\Delta^{1+\beta}) \right)
\]

(13)
\section{Circle diffeomorphisms}

\subsection{Preparations}

For an orientation-preserving homeomorphism \( T \) of the unit circle \( \mathbb{T}^1 = \mathbb{R}/\mathbb{Z} \), its \emph{rotation number} \( \rho = \rho(T) \) is the value of the limit \( \lim_{i \to \infty} L_T^i(x)/i \) for a lift \( L_T \) of \( T \) from \( \mathbb{T}^1 \) onto \( \mathbb{R} \). It is known since Poincaré that rotation number is always defined (up to an additive integer) and does not depend on the starting point \( x \in \mathbb{R} \). Rotation number \( \rho \) is irrational if and only if \( T \) has no periodic points. We restrict our attention in this paper to this case. The order of points does not depend on the starting point is known since Poincaré that rotation number is always defined (up to an additive integer) and does not depend on the starting point \( x \in \mathbb{R} \). Rotation number \( \rho \) is irrational if and only if \( T \) has no periodic points. We restrict our attention in this paper to this case.

We use the \emph{continued fraction} expansion for the (irrational) rotation number:

\[ \rho = [k_1, k_2, \ldots, k_n, \ldots] = \frac{1}{k_1 + \frac{1}{k_2 + \frac{1}{\ldots + \frac{1}{k_n + \ldots}}}} \in (0, 1) \quad (14) \]

which, as usual, is understood as a limit of the sequence of \emph{rational convergents} \( p_n/q_n = [k_1, k_2, \ldots, k_n] \). The positive integers \( k_n, n \geq 1 \), called \emph{partial quotients}, are defined uniquely for irrational \( \rho \). The mutually prime positive integers \( p_n \) and \( q_n \) satisfy the recurrent relation \( p_n = k_np_{n-1} + p_{n-2}, q_n = k_nq_{n-1} + q_{n-2} \) for \( n \geq 1 \), where it is convenient to define \( p_0 = 0, q_0 = 1 \) and \( p_{-1} = 1, q_{-1} = 0 \).

Given a circle homeomorphism \( T \) with irrational \( \rho \), one may consider a \emph{marked trajectory} (i.e. the trajectory of a marked point) \( \xi_i = T^i\xi_0 \in \mathbb{T}^1, i \geq 0 \), and pick out of it the sequence of the \emph{dynamical convergents} \( \xi_{q_n}, n \geq 0 \), indexed by the denominators of the consecutive rational convergents to \( \rho \). We will also conventionally use \( \xi_{q_{-1}} = \xi_0 - 1 \). The well-understood arithmetical properties of rational convergents and the combinatorial equivalence between \( T \) and \( R_\rho \) imply that the dynamical convergents approach the marked point, alternating their order in the following way:

\[ \xi_{q_{-1}} < \xi_{q_1} < \xi_{q_3} < \cdots < \xi_{q_{2m+1}} < \cdots < \xi_0 < \cdots < \xi_{q_{2m}} < \cdots < \xi_{q_2} < \xi_{q_0} \quad (15) \]

We define the \emph{nth fundamental segment} \( \Delta^{(n)}(\xi) \) as the circle arc \([\xi, T^n\xi]\) if \( n \) is even and \([T^n\xi, \xi]\) if \( n \) is odd. If there is a marked trajectory, then we use the notations \( \Delta^{(n)}_0 = \Delta^{(n)}(\xi_0) \), \( \Delta^{(n)}_i = \Delta^{(n)}(\xi_i) = T^i\Delta^{(n)}_0 \).

The iterates \( T^{qn} \) and \( T^{qn-1} \) restricted to \( \Delta^{(n-1)}_0 \) and \( \Delta^{(n)}_0 \) respectively are nothing else but two continuous components of the first-return map for \( T \) on the segment \( \Delta^{(n-1)}_0 \cup \Delta^{(n)}_0 \) (with its endpoints being identified). The consecutive images of \( \Delta^{(n-1)}_0 \) and \( \Delta^{(n)}_0 \) until their return to \( \Delta^{(n-1)}_0 \cup \Delta^{(n)}_0 \) cover the whole circle without overlapping (beyond their endpoints), thus forming the \emph{nth dynamical partition}:

\[ \mathcal{P}_n = \{\Delta^{(n-1)}_i, 0 \leq i < q_n\} \cup \{\Delta^{(n)}_i, 0 \leq i < q_{n-1}\} \]
of $\mathbb{T}^1$. The endpoints of the segments from $\mathcal{P}_n$ form the set
\[ \Xi_n = \{ \xi_i, 0 \leq i < q_{n-1} + q_n \} \]

Denote by $\Delta_n$ the length of $\Delta^{(n)}(\xi)$ for the rigid rotation $R_\rho$. Obviously enough, $\Delta_n = |q_n \rho - p_n|$. It is well known that $\Delta_n \sim \frac{1}{q_{n+1}}$ (here '$\sim$' means 'comparable', i.e. '$A \sim B$' means $A = \mathcal{O}(B)$ and $B = \mathcal{O}(A')$), thus the Diophantine properties of $\rho \in D_\delta$ can be equivalently expressed in the form:
\[ \Delta_n^{1+\delta} = \mathcal{O}(\Delta_n) \] (16)

We will also have in mind the universal exponential decay property
\[ \frac{\Delta_n}{\Delta_{n-k}} \leq \frac{\sqrt{2}}{(\sqrt{2})^k}, \] (17)
which follows from the obvious estimates $\Delta_n \leq \frac{1}{2} \Delta_{n-2}$ and $\Delta_n < \Delta_{n-1}$.

In [1] it was shown that for any diffeomorphism $T \in C^{3+\beta}(\mathbb{T}^1), T' > 0, \alpha \in [0, 1]$, with irrational rotation number the following Denjoy-type inequality takes place:
\[ (T^{q_n})'(\xi) = 1 + \mathcal{O}(\varepsilon_{n,\alpha}), \quad \text{where} \quad \varepsilon_{n,\alpha} = l_n^{\alpha} + \frac{l_n}{l_{n-1}}q_{n-2}^{\alpha} + \frac{l_n}{l_{n-1}}q_{n-3}^{\alpha} + \cdots + \frac{l_n}{l_0} \] (18)
and $l_m = \max_{\xi \in \mathbb{T}^1} |\Delta_m(\xi)|$. Notice, that this estimate does not require any Diophantine conditions on $\rho(T)$.

Unfortunately, it is not possible to write down a corresponding stronger estimate for $T \in C^{3+\beta}(\mathbb{T}^1), \beta \in [0, 1]$, without additional assumptions. We will assume that the conjugacy is at least $C^\beta$-smooth: $\phi \in C^{1+\gamma}(\mathbb{T}^1), \phi'(\xi) > 0$, with some $\gamma \in [0, 1]$. (Notice, that in conditions of Theorem 1 this assumption holds true with $\gamma = 1 - \delta$ accordingly to [1], and our aim is to raise the value of $\gamma$ to $1 - \delta + \beta$.)

This assumption is equivalent to the following one: an invariant measure generated by $T$ has the positive density $h = \phi' \in C^\gamma(\mathbb{T}^1)$. This density satisfies the homologic equation
\[ h(\xi) = T'(\xi)h(T\xi) \] (19)

The continuity of $h$ immediately implies that $h(\xi) \sim 1$, and therefore $(T^i)'(\xi) = \frac{h(\xi)}{h(T^i\xi)} \sim 1$ and
\[ |\Delta^{(n)}(\xi)| \sim l_n \sim \Delta_n \sim 1 \]
\[ \text{due to } \Delta_n = \int_{\Delta^{(n)}(\xi)} h(\eta) \, d\eta. \]
By this reason, we introduce the notation
\[ E_{n,\sigma} = \sum_{k=0}^{n} \frac{\Delta_n}{\Delta_{n-k}} \Delta_{n-k-1}^{\sigma}, \]
so that $\varepsilon_{n,\alpha}$ in (18) can be replaced by $E_{n,\alpha}$ as soon as we know of the existence of continuous $h$.

It follows also that $(T^i)' \in C^\gamma(\mathbb{T}^1)$ uniformly in $i \in \mathbb{Z}$, i.e.
\[ (T^i)'(\xi) - (T^i)'(\eta) = \mathcal{O}(|\xi - \eta|^\gamma), \] (20)
since \((T^i)'\xi - (T^i)'\eta = \frac{h(\xi)}{h(T\xi)} - \frac{h(\eta)}{h(T\eta)}\) and \(T^i\xi - T^i\eta \sim \xi - \eta\).

The additional smoothness of \(T\) will be used through the following quantities: \(p_n = p_n(\xi_0) = \sum_{i=0}^{q_n-1} \frac{\mathcal{S}T(\xi_i)}{h(\xi_i)} (\xi_i - \xi_{i+q_n-1})\), \(\bar{p}_n = \bar{p}_n(\xi_0) = \sum_{i=0}^{q_n-1} \frac{\mathcal{S}T(\xi_i+q_n)}{h(\xi_i+q_n)} (\xi_i+q_n - \xi_i)\). We have

\[
p_n + \bar{p}_n = \sum_{\xi \in \Xi_n} \mathcal{S}T(\xi) \frac{\hat{\xi} - \xi}{h(\xi)},
\]

where \(\hat{\xi}\) denotes the point from the set \(\Xi_n\) following \(\xi\) in the (circular) order \(\ldots \rightarrow \xi_{q_n-1} \rightarrow \xi_0 \rightarrow \xi_{q_n} \rightarrow \ldots\). It is easy to see that \(N_n(\xi_i) = \xi_{i+q_n}\) for \(0 \leq i < q_n-1\) and \(N_n(\xi_i) = \xi_{i-q_n-1}\) for \(q_n-1 \leq i < q_n + q_n-1\).

In the next two subsections, we will establish certain dependencies between the Denjoy-type estimates in the forms \((T^{q_n})'(\xi) = 1 + \mathcal{O}(\Delta_n^\beta)\) and \((T^{q_n})'(\xi) = 1 + \mathcal{O}(E_n,\sigma)\).

### 3.2 Statements that use the Hölder exponents of \(T^{'''}\) and \(h\)

In all the statements of this subsection, we assume that \(T \in C^{3+\beta}\) and \(h \in C^\gamma\), \(\beta, \gamma \in [0,1]\), but do not make any use of Diophantine properties of \(\rho\).

The next lemma corresponds to the exact integral relation \(\int_{\mathbb{T}} \frac{\mathcal{S}T(\xi)}{h(\xi)} d\xi\) first demonstrated in [5].

**Lemma 2.** If \((T^{q_n})'(\xi) = 1 + \mathcal{O}(\Delta_n^\nu)\), then \(p_n + \bar{p}_n = \mathcal{O}(\Delta_{n-1}^{\min\{\beta,2\nu-1\}})\).

**Proof.** Using the representation \(\mathcal{S}T = (\frac{\mathcal{S}T}{T'})' - \frac{1}{2} (\frac{\mathcal{S}T}{T'})^2\), from (21) we derive

\[
p_n + \bar{p}_n = \sum_{\xi \in \Xi_n} \left[ \left( \frac{T''(\hat{\xi})}{T'(\xi)} - \frac{T''(\xi)}{T'(\xi)} \right) \frac{1}{h(\xi)} + \mathcal{O}(\hat{\xi} - \xi)^{1+\beta} \right]
- \frac{1}{2} \sum_{\xi \in \Xi_n} \left( \frac{T''(\xi)}{T'(\xi)} \right)^2 \frac{\hat{\xi} - \xi}{h(\xi)}
= \sum_{\xi \in \Xi_n} \left[ \frac{1}{h(\xi)} - \frac{1}{2} \left( \frac{T''(\xi)}{h(\xi)} \right) \frac{\hat{\xi} - \xi}{h(\xi)} \right] + \mathcal{O}(\Delta_{n-1}^\beta)
\]

Notice that

\[
h(\xi) - h(\hat{\xi}) = \mathcal{O}(|\hat{\xi} - \xi|^{\nu})
\]

due to (19). In particular, (22) implies that the expression in the last square brackets is \(\mathcal{O}(|\hat{\xi} - \xi|^{\gamma})\), hence using the estimate \(T''(\xi) = \frac{T''(\xi) - T''(\hat{\xi})}{\xi - \hat{\xi}} + \mathcal{O}(\hat{\xi} - \xi)\) we get

\[
p_n + \bar{p}_n = \sum_{\xi \in \Xi_n} \left( \frac{T'(\hat{\xi})}{T'(\xi)} - 1 \right) \frac{1}{\xi - \hat{\xi}} \left[ \frac{1}{h(\xi)} - \frac{1}{h(\hat{\xi})} + \frac{1}{2} \left( \frac{T'(\hat{\xi})}{T'(\xi)} - 1 \right) \frac{1}{h(\hat{\xi})} \right] + \mathcal{O}(\Delta_{n-1}^{\min\{\beta,\nu\}})
\]

Now, the substitutions \(T'(\xi) = \frac{h(\xi)}{h(T\xi)}\) and \(T'(\hat{\xi}) = \frac{h(\xi)}{h(T\xi)}\) transform the last estimate (exactly) into

\[
p_n + \bar{p}_n = \frac{1}{2} \sum_{\xi \in \Xi_n} \frac{h(\xi)}{(h(\xi))^2(\xi - \hat{\xi})} \left[ \left( \frac{h(\xi)}{h(\hat{\xi})} - 1 \right)^2 - \left( \frac{h(T\xi)}{h(T\hat{\xi})} - 1 \right)^2 \right] + \mathcal{O}(\Delta_{n-1}^{\min\{\beta,\nu\}})
\]
Similarly to (22), each one of two expressions in parentheses here are $O(|\hat{\xi} - \xi|^{2\nu})$. It follows, firstly, that

$$p_n + \bar{p}_n = \frac{1}{2} \sum_{\xi \in \Xi_n} \left( \frac{h(T\xi)}{h(T\xi) - 1} \right)^2 \left[ \frac{h(T\hat{\xi})}{(h(T\xi))^2(T\hat{\xi} - T\xi)} - \frac{h(\hat{\xi})}{(h(\xi))^2(\xi - \xi)} \right] + O(\Delta_{n-1}^{\min(\beta, 2\nu - 1)}), \quad (24)$$

since, as it is easy to see, the sums in (23) and in (21) differ by a finite number of terms of the order $O(|\hat{\xi} - \xi|^{2\nu - 1})$, and $2\nu - 1 \leq \nu$. Secondly, we have

$$\frac{h(T\hat{\xi})}{(h(T\xi))^2(T\hat{\xi} - T\xi)} : \frac{h(\hat{\xi})}{(h(\xi))^2(\xi - \xi)} = 1 + \frac{T'(\xi)}{T'(\hat{\xi})} \cdot \left( T'(\xi) : \frac{T\hat{\xi} - T\xi}{\hat{\xi} - \xi} \right) = 1 = O(\hat{\xi} - \xi),$$

so the expressions in the square brackets in (24) are bounded, and therefore the whole sum in it is $\sum_{\xi \in \Xi_n} O(|\hat{\xi} - \xi|^{2\nu}) = O(\Delta_{n-1}^{2\nu - 1})$.

Notice, that Lemma 2 does not use $\gamma$. However, the next one does.

**Lemma 3.** If $(T^{q_n})(\xi) = 1 + O(\Delta_{n}^{\nu})$, then $p_n = O(\Delta_{n-1}^{\min(\beta, 2\nu - 1, \gamma)})$.

**Proof.** It follows from (20) that

$$\frac{|\Delta_i^{(n)}|}{|\Delta_i^{(n-2)}|} : \frac{|\Delta_i^{(n-2)}|}{|\Delta_i^{(n-2)}|} = 1 + O(\Delta_{n-2}^{\gamma}) \quad (25)$$

This implies, together with (22) and $ST(\xi_{i+q_n}) - ST(\xi_i) = O(\Delta_{n}^{\beta})$, that

$$\bar{p}_n + \frac{|\Delta_0^{(n)}|}{|\Delta_0^{(n-2)}|} p_{n-1} = \sum_{i=0}^{q_{n-1} - 1} O(\Delta_n(\Delta_{n-2}^{\gamma} + \Delta_{n}^{\beta} + \Delta_{n}^{\nu})) = \frac{\Delta_n}{\Delta_{n-2}} O(\Delta_{n-2}^{\min(\beta, \gamma, \nu)}) = O(\Delta_{n}^{\min(\beta, \gamma, \nu)})$$

In view of this, Lemma 2 implies $p_n = \frac{|\Delta_0^{(n)}|}{|\Delta_0^{(n-2)}|} p_{n-1} + O(\Delta_{n-1}^{\mu})$, where $\mu = \min\{\beta, 2\nu - 1, \gamma\} \leq 1$. Telescoping the last estimate, we get

$$p_n = \sum_{k=0}^{n} \frac{|\Delta_0^{(n)}| \cdot |\Delta_0^{(n-1)}|}{|\Delta_0^{(n-k)}| \cdot |\Delta_0^{(n-k-1)}|} O(\Delta_{n-k}^{\mu}) = O \left( \Delta_{n-1}^{\mu} \sum_{k=0}^{n} \frac{\Delta_n}{\Delta_{n-k}} \left( \frac{\Delta_{n-1}}{\Delta_{n-k-1}} \right)^{1-\mu} \right),$$

and the latter sum is bounded due to (17).

**Lemma 4.** If $p_n = O(\Delta_{n-1}^{\omega})$, where $\omega \in [0, 1]$, then

$$\text{Dist}(\xi_0, \xi, \xi_{q_{n-1}}, \eta; T^{q_n}) = 1 + (\xi - \eta) O(\Delta_{n-1}^{\min(\beta, \gamma, \omega)}), \quad \xi, \eta \in \Delta_0^{(n-1)};$$

$$\text{Dist}(\xi_0, \xi, \xi_{q_n}, \eta; T^{q_{n-1}}) = 1 + (\xi - \eta) \frac{\Delta_n}{\Delta_{n-2}} O(\Delta_{n-2}^{\min(\beta, \gamma, \omega)}), \quad \xi, \eta \in \Delta_0^{(n-2)}.$$


Proof. Accordingly to (13) and (14), we have

\[
\log \text{Dist}(\xi_0, \xi, \xi_{q_{n-1}}, \eta; T^{q_n}) = \frac{1}{6} \sum_{i=0}^{q_{n-1}} (\xi_i - \xi_{i+q_{n-1}})(T^i \xi - T^i \eta)ST(\xi_i) + (\xi - \eta)O(\Delta_{n-1}^\beta)
\]

On the other hand,

\[
\sum_{i=0}^{q_{n-1}} (\xi_i - \xi_{i+q_{n-1}})(T^i \xi - T^i \eta)ST(\xi_i) - h(\xi_0)(\xi - \eta) p_n
\]

\[
= (\xi - \eta) \sum_{i=0}^{q_{n-1}} (\xi_i - \xi_{i+q_{n-1}})ST(\xi_i) \left[ \frac{T^i \xi - T^i \eta}{\xi - \eta} - (T^i)'(\xi_0) \right] = (\xi - \eta)O(\Delta_{n-1}^\gamma)
\]

because of (20). The first estimate of the lemma follows. To prove the second one, we similarly notice that

\[
\log \text{Dist}(\xi_0, \xi, \xi_{q_n}, \eta; T^{q_{n-1}}) = \frac{1}{6} \sum_{i=0}^{q_{n-1}-1} (\xi_i - \xi_{i+q_n})(T^i \xi - T^i \eta)ST(\xi_i) + (\xi - \eta)O(\Delta_{n-1}^\beta)
\]

and

\[
\sum_{i=0}^{q_{n-1}-1} (\xi_i - \xi_{i+q_n})(T^i \xi - T^i \eta)ST(\xi_i) - h(\xi_0)(\xi - \eta) \frac{|\Delta_0^{(n)}|}{|\Delta_{0}^{(n-1)}|} p_{n-1}
\]

\[
= (\xi - \eta) \sum_{i=0}^{q_{n-1}-1} (\xi_i - \xi_{i+q_n})ST(\xi_i) \left[ \frac{T^i \xi - T^i \eta}{\xi - \eta} - (T^i)'(\xi_0) \right] \frac{|\Delta_0^{(n-2)}|}{|\Delta_0^{(n-1)}|} \frac{|\Delta_0^{(n)}|}{|\Delta_0^{(n)}|}
\]

\[
= (\xi - \eta) \sum_{i=0}^{q_{n-1}-1} (\xi_i - \xi_{i+q_n})ST(\xi_i)O(\Delta_{n-2}^\gamma) = (\xi - \eta) \frac{\Delta_0}{\Delta_{n-2}} O(\Delta_{n-2}^\gamma)
\]

(see (26)).

As in [1], we introduce the functions

\[
M_n(\xi) = D(\xi_0, \xi, \xi_{q_{n-1}}; T^{q_n}), \quad \xi \in \Delta_{0}^{(n-1)};
\]

\[
K_n(\xi) = D(\xi_0, \xi, \xi_{q_n}; T^{q_{n-1}}), \quad \xi \in \Delta_{0}^{(n-2)},
\]

where \(\xi_0\) is arbitrarily fixed. The following three exact relations can be easily checked:

\[
M_n(\xi_0) \cdot M_n(\xi_{q_{n-1}}) = K_n(\xi_0) \cdot K_n(\xi_{q_{n}}), \quad (26)
\]

\[
K_{n+1}(\xi_{q_{n-1}}) - 1 = \frac{|\Delta_0^{(n+1)}|}{|\Delta_{0}^{(n-1)}|} \left( M_n(\xi_{q_{n+1}}) - 1 \right), \quad (27)
\]

\[
\frac{(T^{q_{n+1}})'(\xi_0)}{M_{n+1}(\xi_0)} - 1 = \frac{|\Delta_0^{(n+1)}|}{|\Delta_0^{(n)}|} \left( 1 - \frac{(T^{q_n})'(\xi_0)}{K_{n+1}(\xi_0)} \right), \quad (28)
\]
Proof. Let $O_{1} + \pi \Delta = \pi \Delta$, hence, in fact we have $\nu = 1 + \min\{\beta, \gamma, \omega\}$. This gives us
\[
M_{n}(\xi) = m_{n} + \mathcal{O}\left(\Delta_{n-1}^{\sigma+1}\right), \quad K_{n}(\xi) = m_{n} + \mathcal{O}\left(\Delta_{n}^{\sigma}\Delta_{n-2}\right)
\] (30)
where $m_{n+1}^{2}$ denotes the products in (25). Due to (27) and (30) we have
\[
m_{n+1} - 1 = \left|\frac{\Delta_{0}^{(n+1)}}{\Delta_{0}^{(n-1)}}\right|(m_{n} - 1) + \mathcal{O}(\Delta_{n+1}^{\sigma}\Delta_{n-1}^{\sigma}),
\] (31)
which is telescoped into $m_{n} - 1 = \mathcal{O}(\Delta_{n}E_{n-1,\sigma-1})$, which in turn implies
\[
M_{n}(\xi) = 1 + \mathcal{O}(\Delta_{n-1}E_{n,\sigma-1}), \quad K_{n}(\xi) = 1 + \mathcal{O}(\Delta_{n}E_{n-1,\sigma-1})
\] (32)
(notice that $\Delta_{n-1}E_{n,\sigma-1} = \Delta_{n-1}^{1+\sigma} + \Delta_{n}E_{n-1,\sigma-1}$). Due to (27) and (32) we have
\[
(T_{q}^{n+1})'(\xi_{0}) - 1 = \left|\frac{\Delta_{0}^{(n+1)}}{\Delta_{0}^{(n)}}\right|(1 - (T_{q}^{n})'(\xi_{0})) + \mathcal{O}(\Delta_{n}E_{n+1,\sigma-1})
\] (33)
which is telescoped into
\[
(T_{q}^{n})'(\xi_{0}) - 1 = \mathcal{O}\left(\sum_{k=0}^{n} \Delta_{n-k}^{\sigma}E_{n-k,\sigma-1}\right)
\]
\[
= \mathcal{O}\left(\Delta_{n}^{\sigma} + \sum_{m=0}^{n-k} \Delta_{n-m}^{\sigma}E_{n-m,\sigma-1}\right) = \mathcal{O}\left(\Delta_{n}^{\sigma} + \sum_{s=k}^{n} \Delta_{n-s}^{\sigma}E_{n-s,\sigma-1}\right)
\]
\[
= \mathcal{O}\left(\Delta_{n-s}^{\sigma} + \sum_{s=0}^{n} \Delta_{n}^{\sigma}E_{n-s,\sigma-1}\right) = \mathcal{O}(E_{n,\sigma}),
\]
since $\sum_{k=0}^{n} \Delta_{n-k} = \mathcal{O}(\Delta_{n-s})$ due to (17).}

The summary of this subsection is given by

**Proposition 3.** Suppose that for a diffeomorphism $T \in C^{3+\beta}(\mathbb{T}^{1})$, $T' > 0$, $\beta \in [0, 1]$, with irrational rotation number there exists density $h \in C^{\gamma}(\mathbb{T}^{1})$, $\gamma \in [0, 1]$, of the invariant measure and the following asymptotic estimate holds true: $(T_{q}^{n})'(\xi) = 1 + \mathcal{O}(\Delta_{n}^{\nu})$ with certain real constant $\nu$. Then $(T_{q}^{n})'(\xi) = 1 + \mathcal{O}(E_{n,1+\min\{\beta, \gamma, 2\nu-1\}})$.  

**Proof.** Follows from Lemmas 5 and 3 immediately.

**Remark 3.** In [3] it is shown that for any $T \in C^{3}(\mathbb{T}^{1})$ the following Denjoy-type estimate takes place: $(T_{q}^{n})'(\xi) = 1 + \mathcal{O}(\Delta_{n}^{1/2})$, and in our assumptions it is equivalent to $(T_{q}^{n})'(\xi) = 1 + \mathcal{O}(\Delta_{n}^{1/2})$. Hence, in fact we have $\nu \geq \frac{1}{2}$, though this is of no use for us.
3.3 Statements that use Diophantine properties of $\rho$

Now we start using the assumption $\rho \in D_\delta$, $\delta \geq 0$, however forget about the smoothness of $T$ and the Hoelder condition on $h$.

Lemma 6. If $(T^n)\nu(\xi) = 1 + O(\Delta_n^\nu)$, $\nu \in \left[\frac{\delta}{1+\delta}, 1\right]$, then $h \in C^{\nu(1+\delta)-\delta}(\mathbb{T}^1)$.

Proof. Consider two points $\xi_0, \xi \in \mathbb{T}^1$ and $n \geq 0$ such that $\Delta_n \leq |\phi(\xi) - \phi(\xi_0)| < \Delta_{n-1}$. Let $k$ be the greatest positive integer such that $|\phi(\xi) - \phi(\xi_0)| \geq k\Delta_n$. (It follows that $1 \leq k \leq k_{n+1}$.) Due to the combinatorics of trajectories, continuity of $h$ and the homologic equation (19), we have

$$\log h(\xi) - \log h(\xi_0) = O\left(k\Delta_n^\nu + \sum_{s=n+1}^{+\infty} k_{s+1}\Delta_s^\nu\right),$$

and the same estimate holds for $h(\xi) - h(\xi_0)$, since $\log h(\xi) = O(1)$.

We have $k_{n+1} < \Delta_{n-1}/\Delta_n = O(\Delta_n^{-\frac{\delta}{1+\delta}})$, hence

$$k\Delta_n^\nu = k\nu(1+\delta)^\frac{1+\delta}{\nu} \Delta_n^\nu \cdot k(1+\delta)(1-\nu) \Delta_{n-1}^\nu = O \left( (k\Delta_n)^{\nu(1+\delta)-\delta}\right)$$

and

$$\sum_{m=n+1}^{+\infty} k_{m+1}\varepsilon_m = O \left( \sum_{m=n+1}^{+\infty} \Delta_m^{\nu(1+\delta)-\delta}\right) = O \left( \sum_{m=n+1}^{+\infty} \Delta_{m-1}^{\nu(1+\delta)-\delta}\right) = O \left( \Delta_n^{\nu(1+\delta)-\delta}\right)$$

due to (16) and (17). Finally, we obtain

$$h(\xi) - h(\xi_0) = O((k\Delta_n)^{\nu(1+\delta)-\delta}) = O(|\phi(\xi) - \phi(\xi_0)|^{\nu(1+\delta)-\delta}) = O(|\xi - \xi_0|^{\nu(1+\delta)-\delta})$$

Lemma 7. If $\sigma \in [0, 1+\delta)$, then $E_{n, \sigma} = O(\Delta_n^{\frac{\sigma}{1+\delta}})$.

Proof. Due to (16) we have

$$E_{n, \sigma} = O \left( \Delta_n \sum_{k=0}^{n} \Delta_n^{\frac{\sigma}{1+\delta}-1}\right)$$

The statement of the lemma follows, since $\sum_{k=0}^{n} \Delta_n^{\frac{\sigma}{1+\delta}-1} = O(\Delta_n^{\frac{\sigma}{1+\delta}-1})$ because of (17). \hfill \Box

This subsection is summarized by

Proposition 4. Suppose that for a diffeomorphism $T \in C^1(\mathbb{T}^1)$, $T' > 0$, with rotation number $\rho \in D_\delta$, $\delta \geq 0$, there exists a continuous density $h$ of the invariant measure, and the following asymptotical estimate holds true: $(T^n)^\nu(\xi) = 1 + O(E_{n, \sigma})$ with certain constant $\sigma \in [0, 1+\delta)$. Then $(T^n)^\nu(\xi) = 1 + O(\Delta_n^{\frac{\sigma}{1+\delta}})$ and $h \in C^{\max\{0, \sigma-\delta\}}(\mathbb{T}^1)$.

Proof. Follows from Lemmas 7 and 6 immediately. \hfill \Box
3.4 Proof of Theorem 1

Recall that we need to prove Theorem 1 for \( r = 3 + \beta, \ 0 < \beta < \delta < 1 \). We will use a finite inductive procedure based on Propositions 3 and 4 to improve step by step the Denjoy-type estimate in the form

\[
(T^q_n)'(\xi) = 1 + \mathcal{O}(E_{n,\sigma})
\]  

From [1], it follows that (34) holds true for \( \sigma = 1 \) (see (18)), so this will be our starting point.

Consider the sequence \( \sigma_0 = 1, \sigma_{i+1} = \min \{1 + \beta, \frac{2}{1+\delta} \sigma_i\}, i \geq 0 \). The inductive step is given by the following

Lemma 8. Suppose that \( \sigma_i \in [1, 1 + \beta] \) and (34) holds for \( \sigma = \sigma_i \). Then \( \sigma_{i+1} \in [1, 1 + \beta] \) and (34) holds for \( \sigma = \sigma_{i+1} \).

Proof. First of all, notice that \( \sigma_i < 1 + \delta \) since \( \beta < \delta \). Proposition 4 implies that \( h \in C^{\gamma_i}(\mathbb{T}^1) \) with \( \gamma_i = \sigma_i - \delta \in (0, 1) \) and \( (T^q_n)'(\xi) = 1 + \mathcal{O}(\Delta_{n}) \) with \( \nu_i = \frac{\sigma_i}{1+\delta} \in (0, 1) \). Proposition 3 then implies that (34) holds for \( \sigma = \min \{1 + \beta, 1 + \gamma_i, 2\nu_i\} \), and this is exactly \( \sigma_{i+1} \) since \( 1 + \sigma_i - \delta > \frac{2\sigma_i}{1+\delta} \) (indeed, \( (1 + \sigma_i - \delta)(1 + \delta) - 2\sigma_i = (1 - \delta)(1 + \delta - \sigma_i) > 0 \)). The bounds on \( \sigma_{i+1} \) are easy to derive.

What is left is to notice that \( \sigma_i = \min \{1 + \beta, \left(\frac{2}{1+\delta}\right)^i\}, i \geq 0 \), where \( \frac{2}{1+\delta} > 1 \), so this sequence reaches \( 1 + \beta \) in a finite number of steps. And as soon as (34) with \( \sigma = 1 + \beta \) is shown, Proposition 4 implies that \( h \in C^{1+\beta-\delta} \). Theorem 1 is proven.

Acknowledgement. The author thanks Konstantin Khanin for inspiration and lots of useful discussions.

References

1. K. Khanin and A. Teplinsky. Herman’s theory revisited, arXiv:math.DS/0707.0075.
2. M.-R. Herman. Sur la conjugaison differentiable des difféomorphismes du cercle a des rotations, I. H. E. S. Publ. Math., 49, 5–233 (1979).
3. J.-C. Yoccoz. Conjugaison différentiable des difféomorphismes du cercle dont le nombre de rotation vérifie une condition diophantienne. Ann. Sci. École Norm. Sup. (4), 17:3, 333–359 (1984).
4. Y. Katznelson and D. Ornstein. The differentiability of the conjugation of certain diffeomorphisms of the circle, Ergodic Theory Dynam. Systems, 9:4, 643–680 (1989).
5. Ya. G. Sinai and K. M. Khanin, Smoothness of conjugacies of diffeomorphisms of the circle with rotations, Uspekhi Mat. Nauk 44:1, 57-82 (1989), in Russian; English transl., Russian Math. Surveys 44:1, 69-99 (1989).