Ergod. Th. & Dynam. Sys. (1989), **9**, 681–690
Printed in Great Britain

The absolute continuity of the conjugation of certain diffeomorphisms of the circle

Y. KATZNELSON AND D. ORNSTEIN†
Mathematics Department, Stanford University, Stanford CA 94305, *USA*

(*Received* 22 *June* 1987 *and revised* 27 *August* 1988)

0. *Introduction*

Let f be an orientation preserving $\mathcal{H}^1$-diffeomorphism of the circle. If the rotation number $\alpha=\rho(f)$ is irrational and $\log Df$ is of bounded variation then, by a well-known theorem of Denjoy, f is conjugate to the rigid rotation R_α. The conjugation means that there exists an essentially unique homeomorphism h of the circle such that $f=h^{-1}R_\alpha h$. The general problem of relating the smoothness of h to that of f under suitable diophantine conditions on α has been studied extensively (cf. [$\mathbf{H_1}$], [**KO**], [**Y**] and the references given there). At the bottom of the scale of smoothness for f there is a theorem of M. Herman [$\mathbf{H_2}$] which states that if Df is absolutely continuous and $D\log Df\in L^p$, $p>1$, $\alpha=\rho(f)$ is of 'constant type' which means 'the coefficients in the continued fraction expansion of α are bounded', and if f is a perturbation of R_α, then h is absolutely continuous. Our purpose in this paper is to give a different proof and an improved version of Herman's theorem. The main difference in the result is that we do not need to assume that f is close to R_α; the proof is very different from Herman's and is very much in the spirit of [**KO**].

It is not hard to see that the condition of boundedness of the continued-fraction coefficients of α is essential. Given α with unbounded coefficients one can construct $f\in\mathcal{H}^2$ such that h is purely singular (see e.g., [**HS**], [**K**], [**L**]).

This paper assumes a general understanding of the dynamics of circle rotations. We shall refer to [**KO**] for some of the basic facts and notations (but not to the main results of [**KO**] which assume more smoothness of f and give more for h).

1. *Notation, terminology and some background*

Our setup is as follows: $f=h^{-1}R_\alpha h$ is a diffeomorphism of the circle $\mathbf{T}=\mathbf{R}/\mathbf{Z}$, h is a homeomorphism and R_α is the rigid rotation by α. We assume that α, which is defined mod 1, is irrational and, taking a representative in $(0,1)$ we denote by a_n the coefficients of the continued fraction expansion of α, so that

$$\alpha=\cfrac{1}{a_1+\cfrac{1}{a_2+\cdots}}$$

† Research partially supported by NSF Grant No. DMS86-05098.

and the denominators q_n of the convergents satisfy† $q_{n+1} = a_n q_n + q_{n-1}$.

Definition 1.1. An interval $I = (t, \tau)$ is *q_n-small* and its endpoints t, τ are *q_n-close* if $\{f^j(I)\}_{j=0}^{q_n-1}$ are disjoint.

One checks easily that (t, τ) is q_n-small if, depending on the parity of n, either $t \le \tau \le f^{q_{n-1}}(t)$ or $f^{q_{n-1}}(\tau) \le t \le \tau$. The following simple observation is a convenient starting point for 'the basic procedure' (see e.g. [**KO**]).

LEMMA 1.1. *Let $t, \bar{t} \in \mathbf{T}$ and $n \ge 1$. Then there exist $\tau \in \mathbf{T}$ and an integer l, $0 \le l < q_n$, such that τ is q_n-close to t and $\bar{t} = f^l(\tau)$.*

We assume that $\log Df$ is absolutely continuous and $D \log Df \in L^p$ for some $p > 1$.

Notations:

(a) $\mathbf{K}_n^0 = \|\log Df^{q_n}\|_\infty$

(b) $\tilde{\tilde{\mathbf{K}}}_n^1 = \operatorname{Sup} |\int_t^\tau D \log Df^l(s)\, ds| = \sup |\log Df^l(\tau) - \log Df^l(t)|$ the supremum being taken for all l, $0 \le l < q_n$, and intervals (t, τ) which are q_n-small.

(c) $\bar{\bar{\mathbf{K}}}_{m,n}^1 = \operatorname{Sup} |\int_t^\tau D \log Df^l(s)\, ds|$ the supremum is taken now for l of the form $l = cq_m < q_{m+1}$, $m < n$, and intervals (t, τ) which are q_n-small.

We have the following

LEMMA 1.2.

$$\mathbf{K}_n^0 \le 2\tilde{\tilde{\mathbf{K}}}_n^1, \tag{1.1}$$

$$\tilde{\tilde{\mathbf{K}}}_n^1 \le \sum_{m=1}^{n-1} \bar{\bar{\mathbf{K}}}_{m,n}^1. \tag{1.2}$$

Proof. (1.2) follows immediately upon writing an arbitrary l in $(0, q_n)$ as $\sum_m c_m q_m$ with $c_m q_m < q_{m+1}$. (1.1) is closely related to Denjoy's original inequality: one uses the fact that for some $\bar{t} \in \mathbf{T}$, $\log Df^{q_n}(\bar{t}) = 0$, apply lemma 1.1 to obtain τ and l as described there, write

$$\log Df^{q_n}(t) = (\log Df^{q_n}(t) - \log Df^{q_n}(\tau)) + (\log Df^l(\tau) - \log Df^l(f^{q_n}(\tau)))$$

and both differences are bounded by $\tilde{\tilde{\mathbf{K}}}_n^1$. □

We denote, for $m < n$,

$$\eta_n(t) = |f^{q_{n-1}}(t) - t|, \quad \eta_n = \|\eta_n(t)\|_\infty,$$
$$\eta_{m,n}(t) = \eta_n(t)/\eta_m(t), \quad \eta_{m,n} = \|\eta_{m,n}(t)\|_\infty, \tag{1.3}$$

and note that (see e.g., [**KO**] Lemma A.1.1) that there exists $\eta < 1$ which depends only on $\operatorname{Var}(\log Df)$ such that for $n - m \ge 2$

$$\eta_{m,n} \le \eta^{n-m}. \tag{1.4}$$

LEMMA 1.3. *If I is q_n-small, $m < n$, then, with $\mu = dt$ (the Haar measure of $\mathbf{T}$),*

$$\mu\left(\bigcup_{j=0}^{q_{m+1}-1} f^j(I)\right) \le \eta^{n-m}.$$

† Notice that the coefficient which we denote by a_n is denoted by most authors by a_{n+1} (so that, in their notation, $q_{n+1} = a_{n+1} q_n + q_{n-1}$).

https://doi.org/10.1017/S0143385700005289 Published online by Cambridge University Press

Proof. Let $I' = (t_0, f^{q_m}(t_0))$ be a q_{m+1}-interval which contains I. By (1.4) the relative length of I in I' is $\leq \eta^{n-m}$. The same estimate holds, for the same reason, for the ratio $\mu(f^i(I))/\mu(f^i(I'))$ and the lemma follows from the fact that $\sum \mu(f^i(I')) \leq 1$ since they are disjoint. □

Denote $\eta_1 = \eta^{1-p^{-1}}$.

LEMMA 1.4. $\bar{\bar{\mathbf{K}}}^1_{m,n} = O(\eta_1^{n-m})$.

Proof. $\bar{\bar{\mathbf{K}}}^1_{m,n}$ is the integral of $D \log Df$ on a set $U = \bigcup_{j=0}^{c_m q_m - 1} f^j(I)$ and by Lemma 1.3 $\mu(U) \leq \eta^{n-m}$. That means, with $p'^{-1} + p^{-1} = 1$,

$$\|1_U\|_{p'} \leq \eta_1^{n-m} \tag{1.5}$$

and

$$\bar{\bar{\mathbf{K}}}^1_{m,n} = \left|\int_U D \log Df\, dt\right| = \left|\int 1_U D \log Df\, dt\right| \leq \|1_U\|_{p'} \|D \log Df\|_p \tag{1.6}$$

which proves the lemma. □

In the same way we prove that if t, τ are q_n-close, $0 \leq l \leq q_m < q_n$, then

$$|\log Df^l(t) - \log Df^l(\tau)| \leq \text{const. } \eta_1^{n-m}. \tag{1.7}$$

As a corollary to Lemma 1.4 we can replace (1.2) by

$$\bar{\bar{\bar{\mathbf{K}}}}^1_n \leq \sum_{m=k}^{n-1} \bar{\bar{\mathbf{K}}}^1_{m,n} + O(\eta_1^{n-k}). \tag{1.8}$$

2. *Condition sufficient for absolute continuity of the conjugation*

Definition 2.1. Two measures μ, ν on the same σ-algebra are L^2*-equivalent* if $\mu = \varphi_1 \nu$ with $\varphi_1 \in L^2(\nu)$ and $\nu = \varphi_2 \mu$ with $\varphi_2 \in L^2(\mu)$.

LEMMA 2.2.† *Let g be monotone increasing on* $[0, 1]$ *with* $g(0) = 0$, $g(1) = 1$. *Assume that for some sequence* $\{b_n\}$ *such that* $\sum b_n^2 < \infty$ *we have*‡

$$\left|\frac{g(s+2^{-n}) - g(s)}{g(s) - g(s-2^{-n})} - 1\right| < b_n \quad \text{for } 2^{-n} \leq s \leq 1 - 2^{-n}. \tag{2.1}$$

Then g and g^{-1} are absolutely continuous with square-summable derivatives.

Proof. Denote by G_n the linear interpolations of g off $\{j2^{-n}\}_{j=0}^{2^n}$. Then $\{DG_n\}$ is a martingale (relative to the partitions determined by $\{j2^{-n}\}_{j=0}^{2^n}$, $n = 1, 2, \ldots$) and DG_n^{-1} is a martingale relative to the g-image partitions.

Condition (2.1), for $s = (2j+1)2^{-n}$, implies that

$$|\psi_n| = |DG_n - DG_{n-1}| \leq b_n DG_{n-1}$$

and since $\psi_n \perp DG_{n-1}$, we have

$$\|DG_n\|^2_{L^2} = \|\psi_n\|^2_{L^2} + \|DG_{n-1}\|^2_{L^2} \leq (1 + b_n^2)\|DG_{n-1}\|^2_{L^2}$$

and $\|DG_n\|^2_{L^2} \leq \prod_1^n (1 + b_j^2)$. It follows that DG_n converges in L_2 (to Dg).

† An almost identical result appears in [C].
‡ We only need (2.1) for s of the form $(2j+1)2^{-n}$, $j = 0, \ldots, 2^{n-1} - 1$.

https://doi.org/10.1017/S0143385700005289 Published online by Cambridge University Press

The geometric meaning of $|\psi_n| \le b_n DG_n$ is that every slope that we see in G_{n-1} is replaced in G_n by two slopes, one bigger and one smaller, but the ratios of the new slopes to the preceding lie in $(1-b_n, 1+b_n)$. If we look now at the inverse mapping, all the slopes are replaced by their reciprocals and the ratios are now bounded by $((1+b_n)^{-1}, (1-b_n)^{-1})$ which is as good as above, and we conclude $Dg^{-1} \in L^2$, as we did for Dg. □

Remark. The condition $\sum b_n^2 < \infty$ is sharp: given a sequence $\{b_n\}$, $b_n > 0$, such that $\sum b_n^2 = \infty$, one can construct a singular g satisfying (2.1) (cf, **[C]**).

Recall that h denotes the homeomorphism which conjugates f with R_α, and dh is the f-invariant measure on $\mathbf{T}$.

THEOREM 2.3. *Assume $\sum (a_n \mathbf{K}_n^0)^2 < \infty$. Then dh and dt are L^2-equivalent.*

Proof. Without loss of generality we may assume $h(0) = 0$ so that $\tilde{h}$, the lifting of h to $\mathbf{R}$, and $\tilde{h}^{-1}$ map $[0, 1]$ onto $[0, 1]$. We want to apply Lemma 2.2 with $g = \tilde{h}^{-1}$, and we just need to show that the assumption $\sum (a_n \mathbf{K}_n^0)^2 < \infty$ implies (2.1) with $\sum b_n^2 < \infty$.

Fix n. Take an interval $[t, r] = [s - 2^{-n}, s + 2^{-n}]$ and denote its $\tilde{h}$-preimage by $[\tau, \rho]$ and the $\tilde{h}$-preimage of the midpoint s by σ. We are looking for an estimate b_n for $|(\rho - \sigma)/(\sigma - \tau) - 1|$, and obtain it through an algorithm to find σ using powers of f. We use the notation $d_m = \|q_m \alpha\|$ (the distance of $q_m \alpha$ to the nearest integer on $\mathbf{R}$ or to zero on $\mathbf{T}$) and the relation $a_m = [d_{m-1}/d_m]$. Denote by l the smallest integer such that $d_l < 2^{1-n} = t - r$, and put $c_l = [2^{1-n}/d_l]$. Since $2^{1-n} \le d_{l-1}$, we have $c_l \le a_l$. Write $t_1 = t$, $r_1 = r$, $t_2 = r_1 - c_l d_l$, $r_2 = t_1 + c_l d_l$, and observe that $[t_1, t_2]$ is mapped onto $[r_2, r_1]$ by a translation to the right by $c_l d_l$ which is the same as $R_{\pm c_l q_l \alpha}$, (the sign depending on the parity of l). Thus (t_1, r_1) and (t_2, r_2) are concentric and $r_2 - t_2 < c_l d_l$. We now repeat the process for (t_2, r_2): the index l may have increased or remained the same, however, if l remains, that is, $r_2 - t_2 > d_l$, the parameter c_l is certainly lower. Thus we obtain two sequences $\{t_j\}$ and $\{r_j\}$ such that $t_{j+1} > t_j$ and $r_{j+1} < r$, and the interval (t_j, t_{j+1}) is mapped onto (r_{j+1}, r_j) by translation to the right by $c_{l,j} d_{l(j)}$, that is, by $R_{\pm c_{l,j} q_{l(j)} \alpha}$ with $l(j)$ monotone non-decreasing function of j, $c_{l,j} \le a_{l(j)}$ and is (strictly) decreasing on every j-interval on which $l(j)$ is constant. Finally, $r_{j+1} - t_{j+1} < c_{l,j} d_{l(j)}$.

The entire scheme, with $t_1 = t$ and $r_1 = r$ is transported by $\tilde{h}^{-1}$ and gives the sequences $\{\tau_i\}$ and $\{\rho_j\}$ the first increasing to σ, the second decreasing to it, and $[\tau_j, \tau_{j+1}]$ *is mapped onto* $[\rho_{j+1}, \rho_j]$ by $f^{\pm c_{l,j} q_{l(j)}}$. This gives the estimate

$$\left|\frac{\rho_j - \rho_{j+1}}{\tau_{j+1} - \tau_j}\right|^{\pm 1} \le \exp(c_{l,j} \mathbf{K}_{l(j)}^0) \le \exp(a_{l(j)} \mathbf{K}_{l(j)}^0). \tag{2.2}$$

Combining the estimates (2.2) for all j, we obtain

$$\left|\frac{\rho - \sigma}{\sigma - \tau} - 1\right| \le c_l \mathbf{K}_l^0 + \sum_{m=l+1}^{\infty} w_m a_m \mathbf{K}_m^0 = \bar{b}_n + \bar{\bar{b}}_n, \tag{2.3}$$

where w_m are the relative weights of the unions of intervals for which $l(j) = m$. It is not hard to see that $w_m \to 0$ exponentially, in fact since $[\tau, \rho]$ is not q_l-small (though it is q_{l-1}-small), and the interval whose relative measure is denoted by w_m is q_{m-1}-small and is contained in $[\tau, \rho]$, we obtain by (1.4) $w_m \leq \eta^{m-l-1}$.

The manipulation of the rest of the proof is simplest in the case of real interest to us, namely when $a_n = 0(1)$. In this case the parameter l, which is defined by $d_l < 2^{1-n} \leq d_{l-1}$, grows more or less linearly with n (to be precise: any value of l corresponds to at most L values of n or L terms in the martingale, L depends only on the bound for a_n) and the theorem follows from the following (obvious) lemma, putting $\bar{K}_n = a_n \mathbf{K}_n^0$.

LEMMA 2.4. *Assume* $\sum_{n=1}^{\infty} \bar{K}_n^2 < \infty$. *Define* $\bar{b}_n = \sum_{l=n}^{\infty} \eta^{l-n} \bar{K}_l$, *with* $0 < \eta < 1$. *Then* $\sum \bar{b}_n^2 < \infty$.

The proof in the general case follows from the fact that if an interval $n_1 \leq n \leq n_2$ maintains the same value of l, the part $\bar{b}_n (= c_l \mathbf{K}_l^0)$ is largest for $n = n_1$ and drops by a factor $\frac{1}{2}$ as we increase n by one. The part $\bar{\bar{b}}_n = \sum_{l+1}^{\infty} w_m a_m \mathbf{K}_m^0$ is largest for $n = n_2$ and drops by $\frac{1}{2}$ as we decrease n by one. Thus $\sum_{n_1}^{n_2} \leq 2(b_{n_1} + b_{n_2})$ which brings us back to Lemma 2.4 as before. We leave the details to the reader. □

3. *Estimates of* $\|\log Df^{q_n}\|_\infty$

Our main goal here is

THEOREM 3.1. *Assume* $D \log Df \in L^p$, *for some* $p > 1$. *Then*

$$\sum_{n=1}^{\infty} (\mathbf{K}_n^0)^2 < \infty. \tag{3.1}$$

Notice that we do not assume any diophantine condition on α (except, of course, of being irrational). On the other hand, if the coefficients a_n are bounded, then (3.1) implies the condition which, by Theorem 2.3, guarantees the mutual absolute continuity (in fact the L^2-equivalence) of dh and dt.

We shall make use of the following proposition which seems to be in the spirit of Littlewood–Paley, but as far as we know is new.

PROPOSITION 3.2. *Let* $\{G_n\}$ *be an* L^p*-bounded martingale,* $1 < p \leq 2$. *Write* $g_n = G_n - G_{n-1}$. *Then*

$$\sum \|g_n\|_p^2 < \infty. \tag{3.2}$$

We start with

LEMMA 3.3. *Let* V *be a measurable set in a probability space,* $g \in L^p(V)$, $\int_V g\, d\mu = 0$ *and* $\lambda > 0$. *Then (the integrals on the right in* V*):*

$$\int_V (|\lambda + g|^p - \lambda^p)\, d\mu \geq c_p \left(\lambda^{p-2} \int_{|g| < \lambda} g^2\, d\mu + \int_{|g| \geq \lambda} |g|^p\, d\mu \right). \tag{3.3}$$

with $c_p > 0$ *depending only on* p.

Proof. Taylor's theorem with second-order remainder, and direct observation for

$x < -\lambda$ give

$$|\lambda + x|^p - \lambda^p - p\lambda^{p-1}x \geq c_p \begin{cases} \lambda^{p-2}x^2 & \text{for } |x| < \lambda \\ |x|^p & \text{for } |x| \geq \lambda \end{cases} \tag{3.4}$$

and the lemma follows by writing g for x in (3.4) and integrating over V. □

Proof of Proposition 3.2. Set $b_n = \int (|G_n|^p - |G_{n-1}|^p)\, d\mu$ so that $\sum b_n = \sup \|G_n\|_p^p$, and apply† Lemma 3.3 to sets V which are level sets for G_{n-1} to obtain

$$\int_{|g_n| > |G_{n-1}|} |g_n|^p \, d\mu \leq c_p^{-1} b_n \tag{3.5}$$

and

$$\int_{|g_n| < |G_{n-1}|} \left| \frac{g_n}{G_{n-1}} \right|^2 |G_{n-1}|^p \, d\mu < c_p^{-1} b_n \tag{3.6}$$

and since $\int |G_{n-1}|^p \, d\mu$ is bounded and $p \leq 2$ the L^p-norm of g_n/G_{n-1} with respect to the measure $|G_{n-1}|^p \, d\mu$ is bounded by a constant times the L^2-norm and we obtain

$$\left(\int_{|g_n| < |G_{n-1}|} \left| \frac{g_n}{G_{n-1}} \right|^p |G_{n-1}|^p \, d\mu \right)^{1/p} \leq c b_n^{1/2} \tag{3.7}$$

and finally, combining (3.5) and (3.7),

$$\left(\int |g_n|^p \, d\mu \right)^{2/p} \leq c b_n$$

which completes the proof. □

Proof of Theorem 3.1. By (1.1) it is enough to prove

$$\sum (\tilde{\bar{\mathbf{K}}}_n^1)^2 < \infty. \tag{3.1*}$$

We propose to prove (3.1*) by obtaining estimates of $\bar{\bar{\mathbf{K}}}_{m,n}^1$ and then, invoke Lemma 1.2 and (1.8).

So let $n > 0$ be arbitrary, $m < n$ (by (1.8) we shall need only consider $n/2 < m < n$), $l = c_m q_m < q_{m+1}$ and $I = (t, \tau)$ which is q_n-small. By its definition $\bar{\bar{\mathbf{K}}}_{m,n}$ is the supremum of integrals of the form $|\int_I D \log Df^l(s)\, ds|$ and we now fix values of I and l that give the supremum. Keeping in mind that

$$\int_I D \log Df^l(s)\, ds = \int_U D \log Df(s)\, ds$$

with $U = \bigcup_{j=0}^{l-1} f^j(I)$, and we can then rewrite U as $\bigcup_{j=0}^{q_m-1} f^j(E)$ with $E = \bigcup_{k=0}^{c_m-1} f^{kq_m}(I)$. Notice that the condition $c_m q_m < q_{m+1}$ implies that E is contained in a q_m-interval. We now look for a q_m-interval $J = (\bar{t}, f^{q_{m-1}}(\bar{t}))$ such that, writing $V = \bigcup_{i=0}^{q_m-1} f^i(J)$, we have $\int_V D \log Df\, dt = 0$, (we obtain it by noting that the integral is equal to $\log Df^{q_m}(f^{q_{m-1}}(\bar{t})) - \log Df^{q_m}(\bar{t})$ which is continuous in $\bar{t}$, has mean value zero (relative to dh) and must therefore change signs). The measure $\mu(V)$ is clearly bounded by 1 but, as $V \cup f^{q_m}(V)$ is the entire circle and Df^{q_m} is uniformly

† The martingale condition $E(g_n \mid G_{n-1}) = 0$ supplies the needed $\int_V g\, d\mu = 0$.

bounded ($\leq \exp(\text{Var} \log Df)$ by Denjoy's inequality) we obtain also a lower bound and hence by Lemma 1.3 we have

$$\rho_{m,n} = \mu(U)/\mu(V) \leq C\eta^{n-m},$$

and writing

$$\Phi_{m,n} = 1_U - \rho_{m,n} 1_V$$

we have $\int \Phi_{m,n}(t)\,dt = 0$.

The whole idea of the basic procedure is to evaluate a sum, here taking the form of an integral on a set U, by comparing it to one of the same form which is known to vanish. Thus, instead of evaluating $\int 1_U D \log Df\,ds$ we evaluate $\int \Phi_{m,n} D \log Df\,ds$.

By Lemma 1.1 there exist τ which is q_n-close to a point in I and such that $\bar{t} = f^l(\tau)$ with $0 \leq l < q_m$. We write $J^* = f^{-l}(J)$ and

$$j^* = \begin{cases} j & l \leq j < q_m \\ j + q_m & 0 \leq j < l, \end{cases} \tag{3.8}$$

so that $V = \bigcup_{j=0}^{q_m - 1} f^{j^*}(J^*)$. The advantage of this notation is that $f^{j^*}(J^*)$ is q_m-close to $f^j(E)$ (i.e., some points in the one are q_m-close to points in the other.)

One can compare $\mu(f^j(E)$ and $\rho_{m,n}\mu(f^{j^*}(J^*))$ by noticing first that on the average they are equal, which implies that $\mu(f^j(E)) \geq \rho_{m,n}\mu(f^{j^*}(J^*))$ for some values of j, while the opposite (non-strict) inequality holds for other (values of j). On the other hand, for any $t_1 \in f^{j_1}(E)$ and $t_2 \in f^{j_1^*}(J^*)$, as t_1 and t_2 are either q_m-close or at worst both are q_m-close to some t_3; and as for any j_2 in our range we have $j_2^* - j_1^* = j_2 - j_1 + \varepsilon q_m$ with $\varepsilon = \pm 1$ or zero, we obtain (invoking (1.1) if $\varepsilon \neq 0$)

$$|\log Df^{j_2 - j_1}(t_1) - \log Df^{j_2^* - j_1^*}(t_2)| < 4\tilde{\tilde{\mathbf{K}}}_m^1 \tag{3.9}$$

which implies

$$|\log(\mu(f^{j_2}(E))/\mu(f^{j_1}(E))) - \log(\mu(f^{j_2^*}(J^*))/\mu(f^{j_1^*}(J^*)))| \leq 4\tilde{\tilde{\mathbf{K}}}_m^1$$

and, since for any $j = j_2$ we can find j_1 such that the signs of

$$\mu(f^{j_1}(E)) - \rho_{m,n}\mu(f^{j_1^*}(J^*)) \quad \text{and} \quad \mu(f^{j_2}(E)) - \rho_{m,n}\mu(f^{j_2^*}(J^*))$$

are opposite, we obtain

$$|\log \mu(f^j(E)) - \log[\rho_{m,n}\mu(f^{j^*}(J^*))]| \leq 4\tilde{\tilde{\mathbf{K}}}_m^1. \tag{3.10}$$

Define γ_j, $\tilde{\Phi}_{m,n}$ and $\bar{\Phi}_{m,n}$ successively by

$$\begin{aligned} \rho_{m,n}(1+\gamma_j) &= \mu(f^j(E))/\mu(f^{j^*}(J^*)), \\ \tilde{\Phi}_{m,n} &= \rho_{m,n} \sum \gamma_j 1_{(f^{j^*}(J^*))}, \\ \bar{\Phi}_{m,n} &= \Phi_{m,n} + \tilde{\Phi}_{m,n}. \end{aligned} \tag{3.11}$$

Notice that the choice of γ_j guarantees that

$$\int \bar{\Phi}_{m,n} = 0 \quad \text{on} \quad f^j(E) \cup f^{j^*}(J^*) \tag{3.12}$$

and as $\tilde{\tilde{\mathbf{K}}}_m^1 \to 0$, (3.10) implies that (for $m > m_0$)

$$|\gamma_j| \leq 4\tilde{\tilde{\mathbf{K}}}_m^1. \tag{3.13}$$

https://doi.org/10.1017/S0143385700005289 Published online by Cambridge University Press

Remember that we are trying to evaluate $\bar{\bar{\mathbf{K}}}^1_{m,n}$ which is now given as

$$\bar{\bar{\mathbf{K}}}^1_{m,n} = \left| \int \Phi_{m,n} D \log Df\, ds \right| = \left| \int (\bar{\Phi}_{m,n} - \tilde{\Phi}_{m,n}) D \log Df\, ds \right| \tag{3.14}$$

and we estimate separately $\int \bar{\Phi}_{m,n} D \log Df\, ds$ and $\int \tilde{\Phi}_{m,n} D \log Df\, ds$. For the latter we need to point out not only that ((3.13))

$$\|\tilde{\Phi}_{m,n}\|_\infty \le 4\rho_{m,n} \tilde{\bar{\mathbf{K}}}^1_m, \tag{3.15}$$

but also that γ_j changes very slowly with j. Specifically, if we fix b and impose $|j_1 - j_2| < q_b$ then, by Lemma 1.4,

$$|\gamma_{j_1} - \gamma_{j_2}| = O(\eta^{m-b}). \tag{3.16}$$

For sufficiently large b we write $B = \bigcup f^{kq_b}(J)$, $kq_b < q_m$, and by (3.16)

$$\int \tilde{\Phi}_{m,n} D \log Df\, ds = \int_B \tilde{\Phi}_{m,n} D \log Df^{q_b}\, ds + O(\eta^{m-b})\rho_{m,n}. \tag{3.17}$$

B is contained in a q_b-small interval and we can invoke [**KO**] Theorem 3.9 and (3.15) to obtain for $m > m(\varepsilon)$,

$$\left| \int \tilde{\Phi}_{m,n} D \log Df\, ds \right| \le \varepsilon \rho_{m,n} \tilde{\bar{\mathbf{K}}}^1_m + O(\eta^{m-b})\rho_{m,n}, \tag{3.18}$$

where we may take ε arbitrarily small, (determine b by Theorem 3.9 of [**KO**] and take $m > b$). The only thing we shall want from ε is to be small enough (less than some constant that we specify later) we can fix it as well as b once and for all, absorb the factor η^{-b} into the constant, and remembering that $\rho_{m,n} \le C_* \eta^{n-m}$ (3.18) becomes

$$\left| \int \tilde{\Phi}_{m,n} D \log Df\, ds \right| \le \varepsilon C_* \eta^{n-m} \tilde{\bar{\mathbf{K}}}^1_m + O(\eta^n). \tag{3.19}$$

For the estimate of $\int \bar{\Phi}_{m,n} D \log Df\, ds$ we denote by $P_j = P_j(f)$ the partition of the circle determined by the points $\{f^i(0)\}_{i=0}^{q_n-1}$, and by $\{G_j\}$ the martingale expansion of $D \log Df$ relative to $\{P_j\}$ (that means that on each interval-atom of P_j, G_j is equal to the mean value of $D \log Df$ on that interval). We write $g_j = G_j - G_{j-1}$ and keep in mind that g_j has integral zero on every $P_{j-1}(f)$-interval. As $\|G_j\|_p \le \|D \log Df\|_p$ we may apply Proposition 3.2 and conclude that for our specific $\{g_j\}$, (3.2) is valid.

We now estimate $\int \bar{\Phi}_{m,n} g_j\, ds$. Both $\bar{\Phi}_{m,n}$ and g_j are simple functions, g_j being measurable P_j and with integral zero on any P_{j-1} atom. Thus, whenever $\bar{\Phi}_{m,n}$ is constant on a P_{j-1} atom we get no contribution from that atom to $\int \bar{\Phi}_{m,n} g_j\, ds$. Similarly, when $f^k(E) \cup f^{k^*}(J^*)$ is contained in a P_j atom (or, more generally, when g_j is constant on $f^k(E) \cup f^{k^*}(J^*)$) we invoke (3.12) and again get zero contribution to the integral. As we verify below, all this implies:

$$\left| \int \bar{\Phi}_{m,n} g_j\, ds \right| \le \begin{cases} C\eta^{n-j} \|g_j\|_1 & j < m \\ C\eta_1^{n-m} \|g_j\|_p & n \ge j \ge m. \\ C\eta_1^{j-m} \|g_j\|_p & j > n \end{cases} \tag{3.20}$$

We check this case by case:

https://doi.org/10.1017/S0143385700005289 Published online by Cambridge University Press

For $j<m$, the contribution to the integral of a given P_j atom happen only when $f^k(E)\cup f^{k^*}(J^*)$ is partly in the atom but not completely, which happens for two values of k at most. $f^k(E)$ has relative measure in the atom bounded by η^{n-j}, $f^{k^*}(J^*)$ has its relative measure bounded by η^{m-j} and $\Phi_{m,n}$ is bounded by $\rho_{m,n}\le C\eta^{n-m}$ on it (outside $f^k(E)$).

For $j\in[m,n]$ the integral on $\bigcup f^k(E)$ is estimated as in the proof of Lemma 1.4; that on $\bigcup f^{k^*}(J^*)$ is (trivially) much smaller.

For $j>n$ one again estimates the measure of the union of the P_{j-1} atoms on which $\bar{\Phi}_{m,n}$ is not constant.

All that we need to do now is put it all together: by (3.14)

$$\bar{\bar{\mathbf{K}}}^1_{m,n}\le\left|\int\bar{\Phi}_{m,n}D\log Df\,ds\right|+\left|\int\tilde{\Phi}_{m,n}D\log Df\,ds\right| \tag{3.21}$$

and we can estimate the first integral by adding up the estimates (3.20) for all j (recall that $D\log Df=\sum g_j$), and the second by (3.19) and obtain

$$\bar{\bar{\mathbf{K}}}^1_{m,n}\le C\left[\sum_{j<m}\eta^{n-j}\|g_j\|_p+\eta_1^{n-m}\sum_{j=m}^{n}\|g_j\|_p+\sum_{j=n+1}^{\infty}\eta_1^{j-m}\|g_j\|_p+\varepsilon\eta^{n-m}\tilde{\tilde{\mathbf{K}}}^1_m+\eta^n\right]$$

with C a constant which depends only on the variation of $\log Df$.

Summing for $m\in[n/2,n]$ we obtain (see (1.8))

$$\tilde{\tilde{\mathbf{K}}}^1_n\le C\varepsilon\sum_{n/2<m<n}\eta^{n-m}\tilde{\tilde{\mathbf{K}}}^1_m+s_n, \tag{3.22}$$

where

$$s_n=\sum_j c_{n,j}\|g_j\|_p+\frac{n}{2}\eta^n$$

with

$$c_{n,j}\le\begin{cases}n\eta^{n/2} & j<n/2\\ \eta_1^{n-j} & n/2\le j\le n\\ \eta_1^{j-n} & n\le j.\end{cases}$$

By the (trivial) inequality

$$\|(b_{n,j})\|\le\sum_k\sup_n|b_{n,n-k}| \tag{3.23}$$

(the norm of the matrix $(b_{n,j})$ is its norm as operator on l^2) applied to $(c_{n,j})$ and by (3.2) we obtain $\sum s_n^2<\infty$.

By (3.23) the matrix R whose entries are $C\varepsilon\eta^{n-m}$ for $n/2<m<n$, and zero elsewhere has norm on l^2 bounded by $2C\varepsilon(1-\eta)^{-1}<\frac{1}{2}$ for ε fixed small enough. By (3.22),

$$(I-R)\{\tilde{\tilde{\mathbf{K}}}^1_n\}\in l^2$$

and multiplying by $(I-R)^{-1}$ we obtain $\{\tilde{\tilde{\mathbf{K}}}^1_n\}\in l^2$. □

REFERENCES

[C] L. Carleson. On mappings, conformal at the boundary. *J. d'Analyse Math.* **19** (1967), 1-13.

[H_1] M. R. Herman. Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations. *Inst. Hautes Études Sci. Publ. Math.* **49** (1979), 5-234.

[H_2] M. R. Herman. Sur les difféomorphismes du cercle de nombre de rotation de type constant. *Conference on Harmonic Analysis in Honor of A. Zygmund*, Vol. II, (1981), 708-725.

[HS] J. Hawkins & K. Schmidt. On C^2-diffeomorphisms which are of type III_1. *Invent. Math.* **66** (1982), 511-518.

[K] Y. Katznelson. The action of diffeomorphism of the circle on the Lebesgue measure. *J. d'Analyse Math.* **36** (1979), 156-166.

[KO] Y. Katznelson & D. Ornstein. The differentiability of the conjugation of certain diffeomorphisms of the circle. *Ergod. Th. & Dynam. Sys.* this issue, preceding paper.

[L] V. F. Lazutkin. Examples of diffeomorphisms of a circle. *Vestnik Leningrad Univ. Math.* **10** (1982), 55-62.

[Y] 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** (1984), 333-359.