PUBLICATIONS MATHÉMATIQUES DE L'I.H.É.S.

LENNART CARLESON

A remark on Denjoy's inequality and Herman's theorem

Publications mathématiques de l'I.H.É.S., tome 49 (1979), p. 235-241
<http://www.numdam.org/item?id=PMIHES_1979__49__235_0>

A REMARK ON DENJOY'S INEQUALITY AND HERMAN'S THEOREM

by LENNART CARLESON

1. In the preceding proof [1] by M. Herman of the Arnold conjecture, the Hurewicz (or the Chacon-Ornstein) ergodic theorem plays an important role and the proof is in this way non-constructive. The purpose of this note is to give a constructive argument which gives a remainder estimate in the basic Denjoy inequality. This argument also makes it possible to avoid the reduction to the case when

$$\int_0^1 |Df(x)|\,dx = V$$

is small, which was used by Herman.

Let us first recall the situation and some basic results from Herman's paper: $f(x)$ is an increasing continuous function on $-\infty < x < \infty$ such that $f(x+1) - f(x) = 1$, $0 < f(0) < 1$ and $f^i(x)$ are the iterates. Sometimes $f(x)$ will be considered on the torus $\mathbf{T}$ (modulo 1) and this will be clear from the context; α is the rotation number, *i.e.*:

$$|f^n(x) - x - n\alpha| < 1\,;$$

α is assumed irrational with continued fraction expansion $[a_1, a_2, \ldots]$ and p_i/q_i are the convergents. There is a homomorphism h of $(0, 1)$ [mod 1], $t = h(x)$, so that

$$h^{-1} \circ f \circ h(t) = t + \alpha.$$

Herman's theorem asserts that if $f(x)$ is smooth, then, for almost all α, it follows that $h(t)$ is also smooth.

There is a unique probability measure μ on $(0, 1)$ which is invariant under $f(x)$ (mod 1) and:

$$\Big|\sum_{i=0}^{q-1} \varphi(f^i(x_0)) - q\int_0^1 \varphi(f(x))\,d\mu(x)\Big| \leq \mathrm{Var}(\varphi)$$

for all denominators $q = q_j$ in the convergents of α. This is Denjoy's inequality.

We shall prove the following theorem—without use of Herman's result but using his ideas:

For almost all α there are constants C *and* β *so that*

$$\Big|\sum_{i=0}^{q_j-1} \varphi(f^i(x_0)) - q_j\int_0^1 \varphi(x)\,d\mu(x)\Big| < Cj^{-\beta}$$

for all φ *on* $\mathbf{T}$ *with* $|\varphi''(x)| \leq 1$. *The same result holds if* $\varphi'(x)$ *only satisfies some Hölder condition.*

Once this is proved it follows that

$$(1.1) \qquad |\log Df^{q_j}| < Cj^{-\beta}$$

and for almost all α:

$$|\log Df^n| \leq C \sum_{j=1}^{c \log n} \frac{a_j}{j^\beta} = O((\log n)^{1-\beta'})$$

with $\beta' < \beta$. This is the crucial estimate needed for Herman's argument. (1.1) also implies the estimate $|f^q(x) - x - p| < q^{\delta - 1}$, $\delta > 0$. See [1], Chapter VIII.

We shall use the letters C, c to denote different constants whose values are immaterial in the context.

2. Let q be one of the q_j and x_0 a fixed point. We define the measure on (0, 1):

$$\nu_q = k \sum_{i=0}^{q-1} Df^i(x_0) \delta_{f^i(x_0)}$$

where k is chosen so that $\nu_q(0, 1) = 1$. Let I be an interval so that $f(\mathrm{I})$ does not contain x_0 or $f^q(x_0)$. Then it is easy to see that

$$\nu_q(f(\mathrm{I})) = f'(\xi)\nu_q(\mathrm{I}) \quad \text{for some} \quad \xi \in \mathrm{I}.$$

If x_0 or $f^q(x_0) \in f(\mathrm{I})$, the situation is a little more complicated.

We first observe that:

$$(2.1) \qquad |f^q(x_0) - x_0| < q^{-\lambda}, \quad \lambda > 0 \ (\mathrm{mod}\ 1)$$

(see [1], VIII, (2.1)) for almost all α since $\sqrt[j]{q_j} \to$ const, almost everywhere. Assume that the length $|\mathrm{I}|$ of I is greater than $q^{-\lambda/2}$, and that *e.g.* $x_0 \in f(\mathrm{I})$. Suppose also that *e.g.* I extends by $\frac{1}{2}|\mathrm{I}|$ to the right of $f^{-1}(x_0)$. For every second q_i, $f^{q_i}(x_0) > x_0$ and by the inequality (2.1):

$$f^{q_i}(x_0) \in f(\mathrm{I}) \quad \text{if} \quad q_i > \sqrt{q}.$$

This is true for at least $c \log q$ different i's, so that

$$\sum_{f^i(x_0) \in f(\mathrm{I})} Df^i(x_0) \geq c \log q$$

since $Df^{q_i}(x_0) \geqq c > 0$, as was observed by Denjoy.

Now:

$$\begin{aligned} \nu_q(f(\mathrm{I})) &= k \sum_{\substack{f^i(x_0) \in f(\mathrm{I}) \\ i = 0, \ldots, q-1}} Df^i(x_0) = kf'(\xi) \sum_{\substack{f^i(x_0) \in \mathrm{I} \\ i = -1, 0, \ldots, q-2}} Df^i(x_0) \\ &= f'(\xi)\nu_q(\mathrm{I}) + O(kf'(\xi)) \\ &= f'(\xi)\left(1 + O\left(\frac{1}{\log q}\right)\right)\nu_q(\mathrm{I}) \end{aligned}$$

since $\nu_q(\mathrm{I}) > ck \log q$. We obtain the following lemma:

Lemma 1. — Let I *be an interval of length* $> q^{-\lambda/2}$. *Then*:

$$\nu_q(f(\mathrm{I})) = f'(\xi)\nu_q(\mathrm{I}), \quad \xi \in \mathrm{I}, \quad \text{if} \quad x_0, f^q(x_0) \notin f(\mathrm{I}),$$

and

$$\nu_q(f(\mathrm{I})) = f'(\xi)\left(1 + O\left(\frac{1}{\log q}\right)\right)\nu_q(\mathrm{I}), \quad \xi \in \mathrm{I},$$

in all cases, provided α *is not in an exceptional set of measure zero.*

3. Next we need some information on the mapping $x = h(t)$. Let ω be an interval on the t-axis and assume that

$$(3.1) \qquad \frac{1}{q_{i-1}} > |\omega| \geq \frac{1}{q_i}, \qquad a = \frac{q_{i+2}}{q_{i-1}}.$$

We bisect ω into two equal intervals ω_1 and ω_2 and we want to estimate $|h(\omega_1)|$ compared to $|h(\omega)|$. From (3.1) follows that

$$\bigcup_{\nu=1}^{2aq_i} f^\nu(h(\omega)) \supset (0, 1)$$

and every point is covered at most $4a^2$ times. A similar statement is true for ω_1. Namely, if $\alpha = \frac{p_{i+1}}{q_{i+1}} + \frac{\delta}{q_{i+1}^2}$, then $\frac{q_{i+1}}{q_{i+2}} < |\delta| < 1$, so that $2aq_i$ iterations of an interval of length q_{i+1}^{-1} gives a complete covering. Furthermore:

$$|f^\nu(h(\omega))| = \left(\prod_{\mu=0}^{\nu-1} f'(\xi_\mu)\right)|h(\omega)|, \qquad \xi_\mu \in f^\mu(h(\omega)),$$

and similarly for ω_1. Hence:

$$\frac{|f^\nu(h(\omega))|}{|f^\nu(h(\omega_1))|} = \frac{|h(\omega)|}{|h(\omega_1)|} \cdot \prod_{\mu=0}^{\nu-1} \frac{f'(\xi_\mu)}{f'(\xi'_\mu)}$$
$$\geq \frac{|h(\omega)|}{|h(\omega_1)|} e^{-ca^2}$$

and

$$\mathrm{C}a^2 \geq \sum_{\nu=1}^{4aq_i} |f^\nu(h(\omega))| \geq \frac{|h(\omega)|}{|h(\omega_1)|} \sum_{\nu=1}^{4aq_i} |f^\nu(h(\omega_1))| e^{-ca^2}$$
$$\geqq e^{-ca^2} \frac{|h(\omega)|}{|h(\omega_1)|}.$$

This gives the following lemma:

Lemma 2. — Let $\frac{1}{q_i} \leq |\omega| < \frac{1}{q_{i-1}}$ *and bisect* ω *into* ω_1 *and* ω_2. *Then, for almost all* α:

$$|h(\omega_1)| \geq \exp\left(\left(-c\frac{q_{i+2}}{q_{i-1}}\right)^2\right)|h(\omega)|$$

and (see (2.1))

$$|h(\omega)| \le |\omega|^\lambda C \frac{q_{i+2}}{q_{i-1}}.$$

4. We shall now describe the exceptional set of α.

Let δ be a small positive number and n a large integer. Denote by B_ℓ the interval:

$$B_\ell(n) : \delta.\ell n \le k < \delta(\ell+1)n$$

$$\delta^{-1}\frac{1}{2}\frac{2^n}{n} \le \ell < \delta^{-1}\frac{3}{4}\frac{2^n}{n}.$$

The intervals $B_\ell(n)$, $n=1, 2, \ldots$ and ℓ as above, are disjoint.

For every B_ℓ, define the number $b_{\ell,n}(\alpha)$:

$$b_{\ell,n}(\alpha) = \operatorname*{Max}_{k \in B_\ell} \frac{q_{k+1}}{q_k}.$$

For fixed (ℓ, n), $b_{\ell,n}(\alpha) \leqq C$ on a set of measure $\geqq 2^{-n\delta}$. For fixed n, $b_{\ell,n}(\alpha) \leqq C$ for $\geqq 2^{3/4n}$ values of ℓ, if we exclude a set E_n of measure $\le 2^{-n}$ and if $\delta < 1/4$. We now do this for all n and consider those α which do not belong to infinitely many E_n. We also exclude those sets of measure zero mentioned earlier.

5. We shall now prove that ν_q converges weakly to Lebesgue measure and shall also obtain an estimate of the error. We first prove that for some suitable $\gamma > 0$ and $C < \infty$:

$$C^{-1} \le \frac{\nu_q(h(\omega))}{|h(\omega)|} \le C, \quad \text{if} \quad |\omega| > q^{-\gamma}. \tag{5.1}$$

Take some q_i so that $\sqrt{q} < q_i < q$ and so that $\frac{q_{i+1}}{q_{i-1}} < C$. This is possible for almost all α. Then:

$$\alpha = \frac{p_i}{q_i} + \frac{\delta_i}{q_i^2}, \quad 1 > \delta_i > c > 0 \text{ (or } < -c).$$

It follows that if $\frac{1}{q_i} < |\omega| < \frac{2}{q_i}$, then $\bigcup_{\nu=1}^{cq_i} f^\nu(h(\omega)) \supset (0, 1)$ and every point is covered a bounded number of times. Since both $\nu_q(I)$ and $|I|$ are transformed by the rules in lemma 1 it follows that:

$$\frac{\nu_q(h(\omega))}{|h(\omega)|}\frac{1}{C} \leqq \frac{\nu_q(f^i(h(\omega))}{|f^i(h(\omega))|} \leqq C\frac{\nu_q(h(\omega))}{|h(\omega)|}$$

and since both measures are additive, (5.1) follows.

We now wish to prove (5.1) with a constant C very close to 1. Let us define M_k by

$$\sup_{|\omega| \ge q_k^{-1}} \frac{\nu_q(h(\omega))}{|h(\omega)|} = M_k.$$

Suppose that $q = q_s$ and choose n so that

$$2^{n-1} < s \leq 2^n.$$

The number of blocks $B_{\ell,n}$ so that M_k increases in $B_{\ell,n}$ by more than a factor $(1 + 2^{-n/2})$ is less than $C.2^{n/2}$. Hence there exists ℓ so that (with $k = \delta \ell n + 2$)

(i) $b_{\ell,n}(\alpha) \leqq C$,

(ii) $M_{k+\delta n} \leq (1 + 2^{-n/2})\, M_k$.

Now pick an interval ω of length between q_k^{-1} and $2q_k^{-1}$, for which

$$\nu_q(h(\omega)) = M_k |h(\omega)|. \tag{5.2}$$

Divide ω into $e^{c\delta n}$ equal intervals ω' by successive bisections. We assert that for every ω':

$$\nu_q(h(\omega')) \geqq M_k |h(\omega')| (1 - e^{-c\delta n}). \tag{5.3}$$

To see this, recall that, by lemma 2,

$$|h(\omega')| \geqq \exp(-c\delta n) |h(\omega)|.$$

Hence if (5.3) is false for one interval ω', it follows by (ii) that

$$\nu_q(h(\omega)) \leqq M_k((1 + 2^{-n/2}) - e^{-c\delta n} . e^{-c\delta n}) |h(\omega)| < M_k |h(\omega)|$$

if δ is small enough. This contradicts the choice (5.2).

Let ω^* be an arbitrary interval of length $|\omega^*|$ so that

$$\frac{1}{20} q_k^{-1} < |\omega^*| < \frac{1}{10} q_k^{-1}.$$

Then for some ω'' of the same length and $\omega'' \subset \omega$:

$$h(\omega^*) = f^m(h(\omega'')), \quad m < C q_k.$$

Divide ω^* and ω'' into intervals of length $e^{-c\delta n} |\omega^*|$ and let ω_0^* and ω_0'' be two corresponding intervals. Then:

$$\nu_q(h(\omega_0^*)) = \prod_{\nu=1}^{m} f'(\xi_\nu)\, \nu_q(h(\omega_0''))(1 + O(2^{-n}))$$

$$|h(\omega_0^*)| = \prod_{\nu=1}^{m} f'(\xi_\nu') |h(\omega_0'')|$$

so that:

$$\frac{\nu_q(h(\omega_0^*))}{|h(\omega_0^*)|} = \frac{\nu_q(h(\omega_0''))}{|h(\omega_0'')|} . \exp\Big(\sum_{\nu=0}^{m} |f^\nu(h(\omega_0''))|\Big)$$
$$= M_k(1 + O(e^{-cn\delta}))(1 + O(e^{-cn\delta}))$$

because, by lemma 2, $|f^\nu(h(\omega_0^*))| \leq e^{-cn\delta} |f^\nu(h(\omega^*))|$, and $\sum_{\nu=0}^{m} |f^\nu(h(\omega^*))| \leq C$.

We cover $(0, 1)$ by disjoint intervals ω^* and obtain:

$$1 = \sum_{\omega^*} \nu_q(h(\omega^*)) = M_k \sum_{\omega^*} |\omega^*| (1 + O(e^{-cn\delta}))$$

so that: $\quad M_k = 1 + O(e^{-cn\delta})$.

Hence, if $|\omega|>q_k^{-1}$, it follows that

$$\nu_q(h(\omega))\leq|h(\omega)|(1+O((\log q)^{-\beta}))$$

and the reverse inequality is proved similarly.

If we observe that $|h(\omega)|\leq(\log q)^{-K}$ for all K if $|\omega|<q^{-c}$, we can conclude that

$$\int_0^1\varphi(x)d\nu_q(x)=\int_0^1\varphi dx+O((\log q)^{-\beta})\quad\text{if}\quad\varphi\in C^1. \tag{5.4}$$

It remains to prove the same remainder estimate in Denjoy's inequality.

We denote by ω_j the interval $\left(\frac{r}{q},\frac{r+1}{q}\right)$ containing $h^{-1}(x_0)+j\alpha$ and denote by ω_{j0} the subinterval $\left(\frac{r+\eta}{q},\frac{r+1-\eta}{q}\right)$ of ω_j where $\eta=q_k/q$ and q_k is the integer defined above. We first observe that

$$\left|\varphi(f^j(x_0))-q\int_{h(\omega_j)}\varphi(x)d\mu(x)\right|<C|h(\omega_j)|.$$

Divide $(0, q-1)$ into blocks $C_1,\dots,C_m$ of length q_k. Since q_k does not divide q we have to skip a set Γ of less than q_k numbers. This set Γ is chosen so that

$$\sum_\Gamma|h(\omega_j)|<\frac{q_k^2}{q}<q^{-c}.$$

To estimate $\sum_{C_\nu}(\varphi(f^j(x_0))-q\int_{h(\omega_j)}\varphi(x)d\mu(x))$ we write $h(\omega_j)=h(\omega_{j0})\cup h(\omega_j\setminus\omega_{j0})$. Then:

$$\sum_\nu\left(\sum_{C_\nu}\left|\varphi(f^j(x_0))2\eta-q\int_{h(\omega_j\setminus\omega_{j0})}\varphi(x)d\mu(x)\right|\right)<C\eta. \tag{5.5}$$

If $C_\nu=(\lambda,\lambda+q_k)$ we set $y_0=f^\lambda(x_0)$. For $y\in h(\omega_{\lambda0})$ we have $j=\lambda+s$,

$$\varphi(f^s(y_0))-\varphi(f^s(y))=\varphi'(z_s)\int_y^{y_0}Df^s(\xi)d\xi$$

and $f^s(y)\in h(\omega_j)$ and z_s is some number in $h(\omega_j)$. Hence:

$$\begin{aligned}\sum_{C_\nu}&\left(\varphi(f^j(x_0))(1-2\eta)-q\int_{h(\omega_{j0})}\varphi(x)d\mu(x)\right)\\&=(1-2\eta)\sum_{s=0}^{q_k-1}\int_{h(\omega_{\lambda0})}\varphi'(z_s)d\mu(y)\int_y^{y_0}Df^s(\xi)d\xi+\text{error}\\&=(1-2\eta)\sum_{s=0}^{q_k-1}\int_{h(\omega_{\lambda0})}d\mu(y)\int_y^{y_0}Df^s(\xi)\varphi'(f^s(\xi))d\xi\\&\quad+O(\operatorname{Max}_\lambda|h(\omega_{\lambda0})|)+\text{error}.\end{aligned}$$

We have used $|\varphi''(x)| \leq 1$. The error occurs because the intervals $h(\omega_{j0})$ are not exactly maps of $h(\omega_{\lambda 0})$. The error is estimated as in (5.5). For the final sum we use (5.4) for $q = q_k$ and find the bound:

$$\sum_{s=0}^{q_k - 1} \int_{h(\omega_{\lambda 0})} d\mu(y) \int_y^{y_0} \mathrm{D}f^s(\xi)\, d\xi \,.\, O((\log q)^{-\beta}) = O\left(\frac{q_k}{q} (\log q)^{-\beta}\right).$$

This proves the remainder estimate.

REFERENCES

[1] HERMAN, Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations, *Publ. Math. I.H.E.S.*, **49** (1979), p. 1-233.

Lennart Carleson
Institut Mittagleffler, Avravagen 17
18262-Djursholm (Sweden)

Manuscrit reçu le 5 août 1978.