The conjugating map for commutative groups of circle diffeomorphisms

Bryna Kra
Hebrew University of Jerusalem

Abstract

For a single aperiodic, orientation preserving diffeomorphism on the circle, all known local results on the differentiability of the conjugating map are also known to be global results. We show that this does not hold for commutative groups of diffeomorphisms. Given a set of rotation numbers, we construct commuting diffeomorphisms in $C^{2-\epsilon}$ for all $\epsilon > 0$ with these rotation numbers that are not conjugate to rotations. On the other hand, we prove that for a commutative subgroup $\mathcal{F} \subset C^{1+\beta}$, $0 < \beta < 1$, containing diffeomorphisms that are perturbations of rotations, a conjugating map h exists as long as the rotation numbers of this subset jointly satisfy a Diophantine condition.

1 Introduction

1.1 Background

Let $\mathcal{H}$ denote the group of orientation preserving homeomorphisms, under composition, of the circle $\mathbf{T}$ and for $\kappa \geq 0$ let $\mathcal{H}^\kappa$ denote the subset of C^κ mappings.

In the 1880's, Poincaré showed that to each such map f is associated a real parameter $\alpha \in [0, 1)$, called the rotation number, and that for aperiodic f the orbit structure of f is the same as that of the rigid rotation by α, $R_\alpha(t) = t + \alpha$ (mod 1). Letting $t \in \mathbf{T}$ and f^j denote the composition $f \circ f \circ \ldots \circ f$ of f with itself j times, the correspondence between an aperiodic element of $\mathcal{H}$ and the associated rotation number means that the two sequences $\{f^j(t)\}_{j \in \mathbf{Z}}$

and $\{j\alpha \pmod 1\}_{j\in\mathbf{Z}}$ can be mapped one onto another by an orientation preserving self mapping of the circle. We use f_α to denote a map in $\mathcal{H}$ with rotation number α and assume from here on that α is irrational, even when not explicitly mentioned. In general we omit the use of the composition symbol.

Half a century later, Denjoy [2] showed that if Df_α is of bounded variation, then f_α is conjugate to the rotation R_α. When f is conjugate to a rotation, the orbit of every point under the iterates of f is dense and f is said to be minimal.

For a commutative subgroup of $\mathcal{H}$, it is easy to show that if one of the diffeomorphisms of the subgroup is conjugate to a rotation, the entire subgroup is simultaneously conjugate to a subgroup of rotations. Thus, joint conjugation of the subgroup is equivalent to the minimality of the subgroup. The differentiability of the conjugation h will be affected, as for a single diffeomorphism, by the differentiability of the diffeomorphisms $f_{\alpha_1}, f_{\alpha_2}, \ldots, f_{\alpha_n}$ and by the joint Diophantine properties of the elements of $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$.

For a single diffeomorphism, all known local results have been shown to be global results: so long as $f \in \mathcal{H}^2$, f is minimal and even if f is a perturbation of a rotation one needs $f \in \mathcal{H}^2$ to ensure minimality. (This problem has an extensive history. See [1], [8], [3], [11], [6], [10] and [5] for results on minimality and differentiability of the conjugating map and [2] and [4] for non-minimal examples.) In contrast, we show that the differentiability required for minimality of commuting diffeomorphisms is different for the local and global situations. In the global situation, the system generated need not be minimal when the diffeomorphisms are less than twice differentiable. However if the diffeomorphisms are sufficiently close to rotations, then the system they generate is minimal. Before stating the theorem precisely, we need some notation.

1.2 Diophantine assumption

Given $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, a finite set of irrational numbers in $[0, 1)$, let $\bar{A}$ denote the finite linear combinations of elements of A,

$$\bar{A} = \left\{ \sum_{j=1}^{n} k_j \alpha_j : k_j \in \mathbf{Z} \text{ for } j = 1, 2, \ldots, n \right\}.$$

Given any $\gamma \in \bar{A}$, define the *length* $l(\gamma)$ of $\gamma = \sum_{j=0}^{n} k_j \alpha_j$ by

$$l(\gamma) = \max_{j=1,2,\ldots,n} |k_j|.$$

For $\gamma \in \mathbf{T}$, define $\|\gamma\|$ to be the distance from γ to the nearest integer, $\|\gamma\| = \min_{k \in \mathbf{Z}} |k - \gamma|$. The set A is said to have *density* $d = d(A)$ if there exists a constant $c > 0$ such that for any $t \in \mathbf{T}$ and for all $m \in \mathbf{N}$, there exists $\gamma \in \bar{A}$ such that

$$\begin{aligned} & l(\gamma) \leq m \\ \text{and} \quad & \|t - \gamma\| \leq c m^{-d}. \end{aligned}$$

(Note that if A has density d, it also has density d' for any $0 < d' < d$.) Thus the density measures how well every point on the circle can be approximated by linear combinations of the α_j, with specific bounds on the sizes of the coefficients.

This is a general condition: almost all n-tuples in $\mathbf{T}$ have density $d > n - \epsilon$ for all $\epsilon > 0$. (See Kra [7] for further details.)

1.3 Statement of results

In Section 2, we show

Theorem 1.1 *Given $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbf{T}$, there exist non-minimal commuting maps $f_{\alpha_1}, f_{\alpha_2}, \ldots, f_{\alpha_n} \in \mathcal{H}^{2-}$.*

(The notation $f \in C^{w-}$ means that $f \in C^{w-\epsilon}$ for all $\epsilon > 0$.)

On the other hand, in Section 3 we show that if the maps are perturbations of rotations we have the existence of a conjugating map, even if none of the α_i individually satisfies a Diophantine condition.

Theorem 1.2 *Let $\mathcal{F} \subset \mathcal{H}$ be a commutative subgroup and $f_{\alpha_1}, f_{\alpha_2}, \ldots, f_{\alpha_n} \in \mathcal{F}$. Assume that $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ has density $d > 1$ and $1/d < \beta < 1$. There exist neighborhoods $N_1, N_2, \ldots, N_n$ of $R_{\alpha_1}, R_{\alpha_2}, \ldots, R_{\alpha_n}$ in $\mathcal{H}^0$ depending on d, such that if $f_{\alpha_j} \in N_j$ for $j = 1, 2, \ldots, n$, then the action of $\mathcal{F}$ is minimal and the conjugation $h \in \mathcal{H}^{1+\beta-1/d-}$.*

See Moser [9] for a local result for n commuting diffeomorphisms of $\mathcal{H}^{\infty}$.

We also prove a result in Section 4 where the commuting diffeomorphisms are assumed to be conjugate in $\mathcal{H}^{\eta}$ to rotations and show that this guarantees that the conjugation is as differentiable as expected by Theorem 1.2. By Theorem 1.1, one must assume some extra condition in order to draw any conclusions.

2 A Denjoy-type example

We show that given irrational numbers $\alpha_1, \alpha_2 \in \mathbf{T}$, there exist commuting diffeomorphisms $f_{\alpha_1}, f_{\alpha_2}$ in $\mathcal{H}^{2-}$ that are not minimal. Although we only consider the case of two diffeomorphisms, the same argument applies for any finite number. The construction is based on examples of Denjoy. (See [2], [4].)

2.1 The semi-conjugating map

Proposition 2.1 *Let $F \subset \mathbf{T}$ be a totally disconnected perfect set and let α_1, α_2 be irrationals in $\mathbf{T}$, independent over the rationals. There exist commuting mappings $f_{\alpha_1}, f_{\alpha_2} \in \mathcal{H}$ such that the limit points of $\{f^j_{\alpha_1}(t)\}_{j\in\mathbf{Z}}$ and $\{f^j_{\alpha_2}(t)\}_{j\in\mathbf{Z}}$, as $j \to \pm\infty$, are exactly F.*

Proof: Enumerate the components $\{I_{n,m}\}_{(n,m)\in\mathbf{Z}^2}$ of $\mathbf{T}\backslash F$ so that their order on $\mathbf{T}$ is the same as that of $\{n\alpha_1 + m\alpha_2\}_{(n,m)\in\mathbf{Z}^2}$. Define f_{α_2} to be linear increasing on $I_{n,m}$ with $f_{\alpha_2}(I_{n,m}) = I_{n,m+1}$ and complete f_{α_2} by continuity. Fixing n, the orbit of any $t \in \mathbf{T}\backslash F$ under f_{α_2} enters each $I_{n,m}$ (for all m) exactly once and so the limit points of $\{f^j_{\alpha_2}(t)\}_{j\in\mathbf{Z}}$ all lie in F. Conversely, any point in F can be written as a limit point of intervals of the form $I_{n,m}$ with fixed n, and so as a limit point of the orbit of t. By construction $\rho(f_{\alpha_2}) = \alpha_2$.

Define f_{α_1} on $\{I_{n,0}\}_{n\in\mathbf{Z}}$ as linear increasing with $f_{\alpha_1}(I_{n,0}) = I_{n+1,0}$. Define f_{α_1} on the remaining intervals by composition. Thus on $I_{n,m}$ for $m > 0$,

$$f_{\alpha_1}(x) = f_{\alpha_2} \cdots f_{\alpha_2}(f_{\alpha_1}(f^{-1}_{\alpha_2} \cdots f^{-1}_{\alpha_2}(x))),$$

where the composition of f_{α_2} (or $f^{-1}_{\alpha_2}$) with itself occurs m times. (Similarly, for $m < 0$, f_{α_1} is defined by composition of $f^{-1}_{\alpha_2}$ and f_{α_2}, each taken m times.)

As for f_{α_2}, the limit points of $\{f^j_{\alpha_1}(t)\}_{j\in\mathbf{Z}}$ as $j \to \pm\infty$ are exactly F and $\rho(f_{\alpha_1}) = \alpha_1$. □

Since $f_{\alpha_1}, f_{\alpha_2}$ commute, the same semi-conjugating map h works for both diffeomorphisms. However by choosing F to be a totally disconnected perfect subset of $\mathbf{T}$, h is not a homeomorphism.

2.2 Differentiability

By taking F to be a set of measure 0 and determining the sizes of the intervals $I_{n,m}$, we can guarantee a greater degree of differentiability for the maps $f_{\alpha_1}, f_{\alpha_2}$. Again, we define f_{α_1} only on $\{I_{n,0}\}_{n\in\mathbf{Z}}$, f_{α_2} on all of $\mathbf{T}$ and then define f_{α_1} elsewhere by composition. On the interval $I_{n,0}$, let

$$\begin{aligned} Df_{\alpha_1} &= 1 \text{ on the endpoints of } I_{n,0} \\ Df_{\alpha_1} &= 1 + 2C_n \text{ at the midpoint of } I_{n,0}, \end{aligned}$$

where C_n is still to be determined and let Df_{α_1} be linear between these values. Then, $|I_{n+1,0}| = |I_{n,0}| + C_n|I_{n,0}|$ and on $\{I_{n,0}\}_{n\in\mathbf{Z}}$, Df_{α_1} has modulus of continuity ψ if and only if $|C_n| = O(\psi(|I_{n,0}|))$. Let

$$\begin{aligned} S &= \sum_{n=-\infty}^{\infty} \frac{1}{(|n|+2)\log^2(|n|+2)} \quad ; \quad s_0 = \frac{1}{2\log^2 S^2} \\ \text{and } |I_{n,0}| &= \frac{s_0}{(|n|+2)\log^2(|n|+2)}. \end{aligned}$$

Then, $|C_n|$ is approximately $\frac{1}{n}$ and so Df_{α_1} has modulus of continuity $\psi(x) = x\log^2 x$ on $\{I_{n,0}\}_{n\in\mathbf{Z}}$.

On the interval $I_{n,m}$, let

$$\begin{aligned} Df_{\alpha_2} &= 1 \text{ on the endpoints of } I_{n,m} \\ Df_{\alpha_2} &= 1 + 2D_m \text{ at the midpoint of } I_{n,m} \end{aligned}$$

and let Df_{α_2} be affine between these values. Then, as above, $|I_{n,m+1}| = |I_{n,m}|(1 + D_m)$ and Df_{α_2} has modulus of continuity ψ if and only if $|D_m| = O(\psi(|I_{n,m}|))$. Letting

$$\begin{aligned} d_n &= \frac{s_0 2\log^2 2}{(|n|+2)\log^2(|n|+2)} \\ \text{and } |I_{n,m}| &= \frac{d_n}{(|m|+2)\log^2(|m|+2)}, \end{aligned}$$

we have, as above, $|D_m|$ is approximately $\frac{1}{m}$ and Df_{α_2} has modulus of continuity $\psi(x) = x\log^2(x)$. Therefore, $Df_{\alpha_2} \in C^{2-}$.

By definition f_{α_1} is the composition of maps, each of which is $\mathcal{H}^{1+\beta}$ for all $\beta < 1$. Since $\mathcal{H}^{1+\beta}$ forms a group, f_{α_1} is itself $\mathcal{H}^{2-}$. Therefore, we have proven Theorem 1.1.

3 The conjugating map in $\mathcal{H}^{2-}$

We consider n commuting diffeomorphisms f_{α_j} which are sufficiently close to the rotations R_{α_j} for $j = 1, 2, \ldots, n$. Before turning to the proof of Theorem 1.2, we clarify the metric used on $\mathcal{H}^{k+\beta}$ and some background material.

3.1 Metric on $\mathcal{H}^{k+\beta}$

It is convenient to use a translation invariant metric on $\mathcal{H}^{k+\beta}$ which is equivalent to the standard metric. Let $\Phi_{k+\beta}(f) = \sup_{\phi\in B_{k+\beta}} \|\phi f\|_{k+\beta}$, where $B_{k+\beta} = \{\phi \in \mathcal{H}^{k+\beta} : \|\phi\|_{k+\beta} = 1\}$ and $\|\phi\|_{k+\beta}$ denotes the standard $\mathcal{H}^{k+\beta}$ norm. Define the $\mathcal{H}^{k+\beta}$ distance of f to the identity I by

$$d_{k+\beta}(f, I) = \log\left(\max(\Phi_{k+\beta}(f), \Phi_{k+\beta}(f^{-1}))\right) + \sup_{t\in\mathbf{T}} |f(t) - t|$$

and the $\mathcal{H}^{k+\beta}$ distance between f and g to be

$$d_{k+\beta}(f, g) = d_{k+\beta}(fg^{-1}, I).$$

3.2 Differentiability and bounds on iterates

The proof uses techniques introduced by Katznelson and Ornstein [5] and exploits the correspondence between the differentiability of the conjugation h and bounds on the iterates of f and its derivatives. Herman [3] gives proofs of the general statement: $h \in \mathcal{H}^{k+\beta}$ if and only if the iterates $\{f^j\}_{j\in\mathbf{Z}}$ are uniformly bounded in $\mathcal{H}^{k+\beta}$. We use a reformulation of these results introduced by Katznelson and Ornstein [5]:

Theorem 3.1 *Let $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be irrationals in $\mathbf{T}$ and let $\bar{A}$ denote the subgroup of rotations generated by A. Assume that $f_\gamma^j = h^{-1}R_\gamma h$ $(\gamma \in \bar{A})$*

are uniformly bounded in $\mathcal{H}^k$, where $k \in \mathbf{N}, k \neq 0$. Assume that the mapping $\gamma \mapsto f_\gamma^j$ from $\{\gamma\}_{\gamma \in \bar{A}}$ into $\mathcal{H}^k$ has modulus of continuity ϕ. Then $h \in \mathcal{H}^k$ and $D^k h$ has modulus of continuity ϕ.

3.3 General approach

In order to prove Theorem 1.2 for general $\mathcal{F} \subset \mathcal{H}$, it suffices consider the properties of a finitely generated subgroup. We plan to show:

Theorem 1.2 *Assume that $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ has density $d > 1$ and $1/d < \beta < 1$. There exist neighborhoods $N_1, N_2, \ldots, N_n$ of $R_{\alpha_1}, R_{\alpha_2}, \ldots, R_{\alpha_n}$ in $\mathcal{H}^0$ depending on d such that if $f_{\alpha_j} \in N_j$ for $j = 1, 2, \ldots, n$, then f_{α_j} is minimal for $j = 1, 2, \ldots, n$ and the conjugation $h \in \mathcal{H}^{1+\beta-1/d-}$.*

Let $\bar{A}$ denote the subgroup of rotations generated by A. Consider the map

$$\gamma \mapsto f_\gamma \tag{1}$$

from the subgroup of the circle $\bar{A}$ to $\mathcal{H}^1$. By Theorem 3.1, the differentiability of h can be determined by estimating the modulus of continuity of this mapping in $\mathcal{H}^1$. Since there exists a translation invariant metric on $\mathcal{H}^1$ (see Section 3.1) it suffices to check the modulus of continuity in a neighborhood of the identity. We start by with some notation.

Let $c = c(d)$ denote the constant determined in the definition of the density of the subgroup. Let

$$E_m = \{\gamma \in \bar{A} : l(\gamma) \leq 2^m\}$$

and

$$E_m^0 = \{\gamma \in E_m : \|\gamma\| \leq c2^{-dm}\}.$$

We check the modulus of continuity of the map in Equation 1 on $\bigcup_m E_m^0$ and show that

$$\|Df_\gamma - 1\|_\infty = O((2^{-dm})^{\beta-1/d-}) \tag{2}$$

for $\gamma \in E_m^0$. As it is easier to work with $\log Df$, we show

$$\|\log Df_\gamma\|_\infty = O((2^{-dm})^{\beta-1/d-}) \tag{3}$$

for $\gamma \in E_m^0$. Since $\log Df_\gamma$ is small, $\|Df_\gamma - 1\|_\infty \sim \|\log Df_\gamma\|$ and Estimate 2 follows from Estimate 3.

Once we have Estimate 3, Df_γ approaches zero exponentially fast in γ for $\beta - 1/d > 0$. Thus, for $\beta - 1/d > 0$ the derivatives Df_γ are uniformly bounded and the map $\gamma \mapsto Df_\gamma$ has modulus of continuity $\beta - 1/d$. By Theorem 3.1, h is differentiable and its derivative has modulus of continuity $\beta - 1/d$.

In order to obtain Estimate 3, we first show that for diffeomorphisms that are ϵ close to rotations, for ϵ sufficiently small, we have this estimate up to a certain m_0. This gives a preliminary estimate for all m, and these estimates can be used to obtain the desired inequality for bigger values of m and eventually all m.

3.4 Proof of the theorem

We start with E_m^0 and proceed by induction. The determination of exactly how large m_0 must be taken is left until later. Fix $\epsilon > 0$, with $\epsilon < \min_{\gamma \in E_m} \|\gamma\|$. Choose neighborhoods $N_1, N_2, \ldots, N_n$ of the rotations $R_{\alpha_1}, R_{\alpha_2}, \ldots, R_{\alpha_n}$ so that for all $\gamma \in E_m$, $|f_\gamma(t) - R_\gamma(t)| \leq \epsilon$ whenever $f_{\alpha_j} \in N_j$ for $j = 1, 2, \ldots, n$. Then for $\gamma \in E_m$,

$$|f_\gamma(t) - t| \leq 2\|\gamma\| \tag{4}$$

for all $t \in \mathbf{T}$. Specifically, for $\gamma \in E_m^0$

$$|f_\gamma(t) - t| \leq 2\|\gamma\| = O(2^{-dm}), \tag{5}$$

where the constant (call it c_m) is a multiple of the density constant c. Fixing m, the neighborhoods $N_1, N_2, \ldots, N_n$, diffeomorphisms $f_{\alpha_1}, f_{\alpha_2}, \ldots, f_{\alpha_n}$ with $f_{\alpha_j} \in N_j$ for $j = 1, 2, \ldots, n$, $\gamma \in E_m^0$ and $x \in \mathbf{T}$, we estimate $\log Df_\gamma(x)$.

For every γ there exists $y = y_\gamma$ such that $\log Df_\gamma(y) = 0$. (Since f_γ maps the circle onto itself, such y exists.) Thus for arbitrary $x \in \mathbf{T}$ we can write $\log Df_\gamma(x) = \log Df_\gamma(x) - \log Df_\gamma(y)$. Since A has density d, we can find $\delta \in E_m$ with $|R_\delta(x) - y| \leq c2^{-dm}$. By choice of the neighborhoods N_j, for $\delta \in E_m$ the map f_δ is close to a rotation and so

$$|f_\delta(x) - y| = O(2^{-dm}) \tag{6}$$

with constant c_m, the same as that in Equation 5.

Write $\gamma = \sum_j \mu_j$, $\delta = \sum_j \nu_j$, where $\mu_j, \nu_j \in A$, and let $\gamma_j = \sum_{i<j} \mu_i$, $\delta_j = \sum_{i<j} \nu_i$. (Thus the last γ_j is γ itself and γ is written as a sum of at most $nl(\gamma) \leq n2^m$ terms. Similarly for δ.) Then:

$$\begin{aligned}
\log Df_\gamma(x) &= \log Df_{-\delta} f_\gamma f_\delta(x) - \log Df_\gamma(y) \\
&= \log Df_\gamma(f_\delta(x)) - \log Df_\gamma(y) + \log Df_\delta(x) \qquad (7) \\
&\quad - \log Df_\delta(f_\gamma(x)) \\
&= \sum_j [\log Df_{\mu_j}(f_{\gamma_j}(f_\delta(x))) - \log Df_{\mu_j}(f_{\gamma_j}(y))] \qquad (8) \\
&\quad + \sum_j [\log Df_{\nu_j}(f_{\delta_j}(x)) - \log Df_{\nu_j}(f_{\delta_j}(f_\gamma(x)))].
\end{aligned}$$

By choice of δ, $|f_\delta(x) - y| = O(2^{-dm})$ (Equation 6) and since $\gamma \in E_m^0$, $|f_\gamma(x) - x| = O(2^{-dm})$ (Equation 5). Furthermore since $\gamma_j \in E_m$

$$\begin{aligned}
|f_{\gamma_j} f_\delta(x) - f_{\gamma_j}(y)| &\leq |f_{\gamma_j} f_\delta(x) - R_{\gamma_j} f_\delta(x)| + |R_{\gamma_j} f_\delta(x) - R_{\gamma_j}(y)| \\
&\quad + |R_{\gamma_j}(y) - f_{\gamma_j}(y)| \\
&\leq c_m 2^{-dm} + |f_\delta(x) - y| + c_m 2^{-dm},
\end{aligned}$$

and so

$$|f_{\gamma_j} f_\delta(x) - f_{\gamma_j}(y)| = O(2^{-dm})$$

with constant at most $3c_m$. Similarly, $|f_{\delta_j} f_\gamma(x) - f_{\delta_j}(x)| = O(2^{-dm})$ with the same constant.

Since $Df_{\alpha_j} \in C^\beta$,

$$|\log Df_{\mu_j}(f_{\gamma_j}(f_\delta(x))) - \log Df_{\mu_j}(f_{\gamma_j}(y))| = O(2^{-dm\beta}) \qquad (9)$$

and

$$|\log Df_{\nu_j}(f_{\delta_j}(x)) - \log Df_{\nu_j}(f_{\delta_j}(f_\gamma(x)))| = O(2^{-dm\beta}) \qquad (10)$$

where the constants depend on c_m and the Hölder constants of $f_{\alpha_1}, f_{\alpha_2}, \ldots, f_{\alpha_n}$. By combining Equations 8, 9 and 10, we have

$$|\log Df_\gamma(x)| = O(nl(\gamma)2^{-dm\beta} + nl(\delta)2^{-dm\beta}) = O((2^{-dm})^{\beta - 1/d}),$$

since $l(\gamma), l(\delta) \leq 2^m$. Therefore, we have the desired conclusion for E_m^0, with some constant c'_m depending on c_m, the Hölder constants of the f_{α_j} and the number of elements n in A.

Next, we show that the same estimate holds for $\gamma \in E^0_{m+1}$. Let $\tilde{\gamma} \in E^0_{m+1}$. Once again, to estimate $\log Df_{\tilde{\gamma}}(x)$, we write $\log Df_{\tilde{\gamma}}(x) = \log Df_{\tilde{\gamma}}(x) - \log Df_{\tilde{\gamma}}(y)$, where $\log Df_{\tilde{\gamma}}(y) = 0$. Since A has density d, there exists $\tilde{\delta} \in E_{m+1}$ with $|R_{\tilde{\delta}}(x) - y| \leq c2^{-d(m+1)}$. As in the expansion of Equation 7,

$$\begin{aligned} |\log Df_{\tilde{\gamma}}(x)| &\leq |\log Df_{\tilde{\gamma}}(f_{\tilde{\delta}}(x)) - \log Df_{\tilde{\gamma}}(y)| \\ &\quad + |\log Df_{\tilde{\gamma}}(x) - \log Df_{\tilde{\delta}}(f_{\tilde{\gamma}}(x))|. \end{aligned} \tag{11}$$

Since $\tilde{\gamma} \in E^0_{m+1}$, $\|\tilde{\gamma}\| \leq 2^{-d(m+1)} \leq 2^{-dm}$. Furthermore, there exists $\gamma \in E_m$ so that $\|\gamma\| \leq 2(2^{-md})$ and $\|\tilde{\gamma}\| \leq \|\gamma\|$. Since the f_{α_j} commute, they are semi-conjugate to rotations and so the order on the circle of the orbit $\{f_\gamma(t)\}_{\gamma\in\bar{A}}$ is the same as that imposed by the rigid rotations $\{R_\gamma(t)\}_{\gamma\in\bar{A}}$. Therefore, by conjugation we have

$$|f_{\tilde{\gamma}}(x) - x| < |f_\gamma(x) - x|,$$

and since the diffeomorphisms are perturbations of rotations, the choice of neighborhoods $N_1, N_2, \ldots, N_n$, gives (see Inequality 4) $|f_\gamma(x) - x| \leq 2c_m 2^{-md}$. Thus,

$$|f_{\tilde{\gamma}}(x) - x| \leq 2c_m 2^{-md}.$$

We apply the same reasoning to $|f_{\tilde{\delta}}(x) - y|$. By conjugation $|f_{\tilde{\delta}}(x) - y|$ is bounded by $|f_\delta(x) - y|$ for some δ, at most twice the distance away, and we already know that $|f_\delta(x) - y| \leq 2c_m 2^{-md}$. Therefore, $|f_{\tilde{\delta}}(x) - y| \leq 2c_m 2^{-md}$.

These estimates can be improved. As for $\gamma \in E_m$, if $\mu \in E_{m+1}$ then $|f_\mu f_{\tilde{\gamma}}(x) - f_\mu(x)| \leq 6c_m 2^{-md}$ and $|f_\mu f_{\tilde{\delta}}(x) - f_\mu(y)| \leq 6c_m 2^{-md}$. By expanding Equation 11 in the same way as Equation 8, combined with the analogous results of 9 and 10, we have

$$\begin{aligned} |\log Df_{\tilde{\gamma}}(x)| &\leq nl(\tilde{\gamma})(3|f_{\tilde{\delta}}(x) - y|)^\beta + nl(\tilde{\delta})(3|f_{\tilde{\gamma}}(x) - x|)^\beta \\ &\leq 2^{m+1} 3c'_m (2^{-md})^\beta + 2^{m+1} 3c'_m (2^{-md})^\beta \\ &= 4(3^\beta c'_m)(2^{-md})^{\beta - 1/d}. \end{aligned} \tag{12}$$

For $\beta - 1/d > 0$, Estimate 12 gives that $|Df_{\tilde{\gamma}}(x)|$ is exponentially close to one. Thus for $m > m_0$ with m_0 sufficiently large the action of $f_{\tilde{\gamma}}$ is like that of rotation by $\tilde{\gamma}$ for a large number of iterations. Considering the interval $[x, f_{\tilde{\gamma}}(x)] \subset [x, f_\gamma(x)]$, where $\gamma \in E_m$ is chosen as above, $f_{\tilde{\gamma}}$ maps the interval $[x, f_{\tilde{\gamma}}(x)]$ onto an adjacent interval. The interval of length γ

(under rotation by $\tilde{\gamma}$) is divided into a certain number of intervals, and by conjugation $f_{\tilde{\gamma}}$ divides the corresponding interval into the same number of pieces. This creates $\|\gamma\|/\|\tilde{\gamma}\|$ subintervals of $[x, f_\gamma(x)]$, each of approximately (exponentially close to being) equal size. That is, we have the improved estimate

$$|f_{\tilde{\gamma}}(x) - x| \leq c'_m 2^{-d(m+1)},$$

where the error in using the constant c'_m is linear in m, and is accounted for by the loss of arbitrarily small ϵ in the conclusion of the theorem.

By the same reasoning, we have $|f_{\tilde{\delta}}(x) - y| \leq c'_m 2^{-d(m+1)}$. Combining these estimates with Equation 12, we have

$$\begin{aligned} |\log Df_{\tilde{\gamma}}(x)| &= O(2^{m+1} 2^{-d(m+1)\beta-}) \\ &= O((2^{-d(m+1)})^{\beta-1/d-}), \end{aligned}$$

where the constant depends on choice of neighborhoods $N_1, N_2, \ldots, N_n$, the number of elements in A, the density constant and the Hölder constants of the diffeomorphisms f_{α_j}. This gives exactly the estimate needed for E_{m+1}.

Since any diffeomorphism that commutes with f_α, when f_α is conjugate to R_α, is also conjugate to a rotation via the same map h, given a commutative subgroup $\mathcal{F} \subset \mathcal{H}$ and some finite subset $f_{\alpha_1}, f_{\alpha_2}, \ldots, f_{\alpha_n} \in \mathcal{F}$ satisfying the hypotheses of Theorem 1.2, we have the existence of $h \in \mathcal{H}^{1+\beta-1/d-}$ that conjugates each $f \in \mathcal{F}$ to a rotation. □

Extensions of these techniques (Kra [7]) can be used to prove local results for higher derivatives, illustrating the exact loss of differentiability that occurs in Moser's Theorem. At this time, a global version of Theorem 1.2 has yet to be proved or disproved.

4 Assuming $\mathcal{H}^\eta$ conjugation

In this section, we prove a result for diffeomorphisms (not necessarily twice differentiable) using a slightly different condition than the local assumption of Theorem 1.2. As the example of Section 2 shows, we need an extra assumption or the conjugating map need not exist. The assumption that h is $\mathcal{H}^\eta$ for some $\eta > 0$ takes the role of the perturbation assumption. We show:

Theorem 4.1 *Let $\mathcal{F} \subset \mathcal{H}^{1+\epsilon}$ be a commuting subgroup with $0 < \epsilon < 1$ and let $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ have density $d > 0$. Assume there exists $h \in \mathcal{H}^\eta$, $0 < \eta < 1$, so that for some $f_{\alpha_1}, f_{\alpha_2}, \ldots, f_{\alpha_n} \in \mathcal{F}$, $f_{\alpha_j} = h^{-1} R_{\alpha_j} h$ for $j = 1, 2, \ldots, n$. Then if $\eta\epsilon - 1/d > 0$, $h \in \mathcal{H}^{1+\eta\epsilon-1/d}$.*

We start with a lemma needed in the proof of the theorem.

Lemma 4.2 *Let $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be elements in $\mathbf{T}$, let $h \in \mathcal{H}$ and $f_{\alpha_j} = h^{-1} R_{\alpha_j} h$ for $j = 1, 2, \ldots, n$. Let $\delta = \sum_{j=1}^n l_j \alpha_j$ and $\gamma = \sum_{j=1}^n k_j \alpha_j$, with $l_j, k_j \in \mathbf{Z}$ for $j = 1, 2, \ldots, n$. For $x, y \in \mathbf{T}$, $\epsilon > 0$ and $0 < \eta < 1$, if $|h(x) + \delta - h(y)| < \epsilon$ then*

$$\begin{array}{ll} (i) & |f_\delta(x) - y| < c\epsilon^\eta \text{ for } h \in \mathcal{H}^\eta \\ & |f_\delta(x) - y| < c\epsilon \text{ for } h^{-1} \in \mathrm{Lip}(1) \\ (ii) & |f_\gamma f_\delta(x) - f_\gamma(y)| < c\epsilon^\eta \text{ for } h \in \mathcal{H}^\eta \\ & |f_\gamma f_\delta(x) - f_\gamma(y)| < c\epsilon \text{ for } h^{-1} \in \mathrm{Lip}(1) \end{array}$$

for some constant c.

Proof: Since h conjugates each diffeomorphism f_{α_j} to a rotation R_{α_j}, given $\delta \in \bar{A}$

$$\begin{aligned} |f_\delta(x) - y| &= |h^{-1}(h(x) + \delta) - h^{-1}(h(y))| \\ &\leq c|h(x) + \delta - h(y)|^\eta \leq c\epsilon^\eta, \end{aligned}$$

when $h^{-1} \in C^\eta$. Similarly, $|f_\gamma f_\delta(x) - f_\gamma(y)| \leq c\epsilon^\eta$ for $h^{-1} \in C^\eta$. For $h^{-1} \in \mathrm{Lip}(1)$, the same argument gives the inequalities with $\eta = 1$. □

We now turn to the theorem. The proof follows the same general steps as the proof of Theorem 1.2.

Proof: Since $\mathcal{F}$ is a commutative subgroup, h simultaneously conjugates each diffeomorphism of $\mathcal{F}$ to a rotation. Thus, it suffices to show that $h \in C^{1+\eta\epsilon-1/d}$. As before, we consider the map

$$\gamma \mapsto f_\gamma = h^{-1} R_\gamma h$$

from $\bar{A}$ to $\mathcal{H}$ and estimate the modulus of continuity in a neighborhood of the identity. Using the same notation as in Theorem 1.2, it suffices to show that for $\gamma \in E_m^0$

$$\|\log Df_\gamma\|_\infty = O((2^{-dm})^{\eta\epsilon-1/d}). \tag{13}$$

Once we have this estimate, the maps Df_γ are uniformly bounded for $\eta\epsilon - 1/d > 0$ and Theorem 3.1 gives $\eta\epsilon - 1/d > 0$, $h \in \mathcal{H}^1$ and Dh has modulus of continuity $\eta\epsilon - 1/d$, and so $h \in \mathcal{H}^{1+\eta\epsilon-1/d}$.

We estimate $\log Df_\gamma(x)$ by writing $\log Df_\gamma(x) = \log Df_\gamma(x) - \log Df_\gamma(y)$, where $\log Df_\gamma(y) = 0$. There exists $\delta \in E_m$ with $|h(x) + \delta - h(y)| < c2^{-dm}$, where c is the constant in the definition of the density of the subgroup. Then:

$$\begin{aligned}\log Df_\gamma(x) &= \log Df_\gamma(f_\delta(x)) - \log Df_\gamma(y) + \log Df_\delta(x) \\ &\quad - \log Df_\delta(f_\gamma(x)).\end{aligned} \tag{14}$$

We estimate the two differences of Equation 14 separately. Writing $\gamma = \sum_j \mu_j$, $\delta = \sum_j \nu_j$ where each $\mu_j, \nu_j \in A$, and setting $\gamma_j = \sum_{i<j} \mu_i$, $\delta_j = \sum_{i<j} \nu_i$, we have for the first term:

$$|\log Df_\gamma(f_\delta(x)) - \log Df_\gamma(y)| \leq \sum_j |\log Df_{\mu_j}(f_{\gamma_j} f_\delta(x)) - \log Df_{\mu_j}(f_{\gamma_j}(y))|.$$

Since $h \in \mathcal{H}^\eta$, we can apply Lemma 4.2. By choice of δ, $|f_\delta(x) - y| \leq c_1 2^{-dm\eta}$, where c_1 is some constant depending on the Hölder constant of h and the density constant. Furthermore, by Lemma 4.2 part (ii), $|f_{\gamma_j}(f_\delta(x)) - f_{\gamma_j}(y)| \leq c_1 2^{-dm\eta}$. Therefore, the sum of the first two terms of Equation 14 is bounded by

$$O(l(\gamma) n 2^{-dm\eta\epsilon}) = O((2^{-dm})^{\eta\epsilon - 1/d}),$$

where the constant depends on the Hölder constants of the f_{α_j}, the number of elements in the subgroup A and the density constant.

Similarly, by applying Lemma 4.2 to the second difference of Equation 14, we have

$$|\log Df_\delta(x) - \log Df_\delta(f_\gamma(x))| = O(l(\delta) n 2^{-dm\eta\epsilon}) = O((2^{-dm})^{\eta\epsilon - 1/d})$$

and again the constant depends on the Hölder constants of the f_{α_j}, the number of elements in A and the density constant. Combining the two terms, we have

$$|\log Df_\gamma(x)| = O((2^{-dm})^{\eta\epsilon - 1/d}),$$

which is exactly Estimate 13. □

Acknowledgments

I would like to thank Y. Katznelson for his invaluable help.

References

[1] V.I. Arnold. Small denominators I, on the mapping of a circle into itself. *Izv. Akad. Nauk. Math.*, 25:21–86, 1961. Transl. A.M.S. Serie 2, 46:213-284, 1965.

[2] A. Denjoy. Sur les courbes défines par les équations différentielles à la surface du tore. *J. Math., Pures et Appl.*, 2:333–375, 1932.

[3] M.R. Herman. Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations. *Publ. Math. I.H.E.S.*, 49:5–234, 1979.

[4] Y. Katznelson. Sigma-finite invariant measures for smooth mappings of the circle. *J. D'Anal. Math.*, 31:1–18, 1977.

[5] Y. Katznelson and D. Ornstein. The differentiability of the conjugation of certain diffeomorphisms of the circle. *Ergod. Th. & Dyna. Sys.*, 9:643–680, 1989.

[6] K.M. Khanin and Ya.G. Sinai. A new proof of M. Herman's theorem. *Commun. Math. Phys.*, 112:89–101, 1987.

[7] B. Kra. *Commutative Groups of Diffeomorphisms of the Circle*. Stanford University, Stanford, 1994.

[8] J. Moser. A rapidly convergent iteration method, part II. *Ann. Scuola Norm. Sup. di Pisa*, 20:499–535, 1966.

[9] J. Moser. On commuting circle mappings and simultaneous diophantine approximations. *Math. Z.*, 205:105–121, 1990.

[10] J. Stark. Smooth conjugacy and renormalization for diffeomorphisms of the circle. *Nonlinearity*, 1:54–575, 1988.

[11] 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):333–359, 1984.