# Commutators of Diffeomorphisms: II.

# Commutators of Diffeomorphisms: II[1])

by John N. Mather

This paper is a sequel to [2], and we will assume the reader is familiar with the terminology and results of [2]. Our main result is the following.

THEOREM 1. *Let $M$ be a smooth $n$-manifold. If $n \geqslant r \geqslant 1$, then Diff $(M, r)$ is perfect.*

From Epstein's theorem [1], it then follows that we have:

COROLLARY. *Under the same hypothesis, Diff $(M, r)$ is simple.*

Combining this with the result in [2], we see that the only missing differentiability class is $r = n + 1$. For $r = n + 1$, it is still unknown whether Theorem 1 holds.

Theorem 1 is an immediate consequence of the case when $M = \mathbf{R}^n$. The proof of Theorem 1 is very closely related to the proof of Theorem 1 in [2]. In a sense, we have turned the proof of the latter upside down.

In the final section, we give a result concerning the connectivity of Haefliger's classifying space as an application of our method. This is analogous to the result concerning the connectivity of Haefliger's classifying space we obtained in [2], but for low differentiability, rather than high differentiability.

## §1. A Refinement of Theorem 1

We will actually prove a refinement of Theorem 1. Let $\alpha$ be a modulus of continuity, and $r$ a positive integer.

THEOREM 2. *Suppose either of the following holds.*

a) $n > r \geqslant 1$

b) $r = n$, $\alpha$ *is defined on all of* $[0, \infty)$, *and there exists $\beta$, with $0 < \beta < 1$, such that* $\alpha(tx) \leqslant t^\beta \alpha(x)$ *for all $x \geqslant 0$ and all $t \geqslant 1$.*

*Then* Diff $(M, r, \alpha)$ *is perfect.*

Theorem 1 is an immediate consequence, since Diff $(M, r) = \bigcup_\alpha$ Diff $(M, r, \alpha)$, where the union is taken over all moduli of continuity in the case $r < n$, and all moduli of continuity satisfying the supplementary condition (b), when $r = n$.

Of course, it is enough to prove Theorem 2 in the case $M = \mathbf{R}^n$, to obtain the

---

general result. We reduce the proof in §2 to the construction of certain mappings $P_{i,A}$. These mappings are constructed in subsequent sections.

## §2. Strategy of the Proof

The main technical step is the construction of certain mappings $P_{i,A}$ of function spaces, and the proof of a number of properties of the $P_{i,A}$. The domain of $P_{i,A}$ is a $C^1$ neighborhood of the identity in the space of $C^1$ diffeomorphisms of $\mathbf{R}^n$ with support in int $D_{i,A}$. The range of $P_{i,A}$ is the set of $C^1$ diffeomorphisms of $\mathbf{R}^n$ with support in int $D_{i-1,A}$.

There is a close parallel between the construction of $P_{i,A}$ which we will give below and the construction of $\Psi_{i,A}$ which we already gave in [2]. Note, however, that there is already a difference: we have reversed the domain and range.

Now we list the properties we will show $P_{i,A}$ to have.

*Properties of $P_{i,A}$.*

1) $P_{i,A}(\text{id}) = \text{id}$,

2) If $u$ is in $C^{r,\alpha}$, then so is $P_{i,A}(u)$.

3) The restriction of $P_{i,A}$ to the set of $C^r$ diffeomorphisms in its domain is continuous with respect to the $C^r$ topologies on its domain and range.

4) If $u$ is in the domain of $P_{i,A}$, then $u$ is isotopic to the identity through an isotopy with support in int $D_{i,A}$ and $P_{i,A}(u)$ is isotopic to the identity through an isotopy with support in int $D_{i-1,A}$.

From 4), if $u \in \text{dom} P_{i,A}$ and $u$ is $C^{r,\alpha}$, then $u, P_{i,A}(u) \in \text{Diff}(\mathbf{R}^n, r, \alpha)$.

5) If $u \in \text{dom} P_{i,A}$ and $u$ is $C^{r,\alpha}$ then $[u] = [P_{i,A}(u)]$ in the commutator quotient group of $\text{Diff}(\mathbf{R}^n, r, \alpha)$.

6) There exists $\delta > 0$, $C > 0$ such that

$$\mu_{r,\alpha}(P_{i,A}(u)) \leqslant CA^{-1}\mu_{r,\alpha}(u),$$

if $u$ is of class $C^{r,\alpha}$, lies in the domain of $P_{i,A}$, and satisfies $\mu_{r,\alpha}(u) < \delta$. Moreover, $C$ is independent of $A$.

The estimate for $P_{i,A}$ given in 6) is in a sense the "inverse" of the estimate for $\Psi_{i,A}$ given in §3, (6) of [2].

In the rest of this section, we finish the proof of the Theorem 2, assuming the existence of $P_{i,A}$ satisfying (1)–(6). Consider $f \in \text{Diff}(\mathbf{R}^n, r, \alpha)$ with support in int $D_n$. We wish to show $f$ is in the commutator subgroup if it is sufficiently close to id.

For any $u \in \text{Diff}(\mathbf{R}^n, r, \alpha)$ with support in int $D_{0,A}$ we try to define

$$u_0 = A^{-1}fuA, \quad u_1 = P_{n,A}(u_0), \quad u_2 = P_{n-1,A}(u_1), \ldots, u_n = P_{1,A}(u_{n-1}).$$

If $u$ and $f$ are sufficiently close to the identity in the $C^1$ topology, these will actually be defined, by properties (1)–(3) of $P_{i,A}$.

It is easily seen that $u_0$ is conjugate to $fu$ in $\mathrm{Diff}(\mathbf{R}^n, r, \alpha)$. Thus, $[u_0] = [fu]$ in the commutator quotient group of $\mathrm{Diff}(\mathbf{R}^n, r, \alpha)$. Then

$$[fu] = [u_n], \tag{$*$}$$

by (5) and the definition of the $u_i$.

LEMMA. *Suppose the hypotheses of Theorem 2 are satisfied. There exists $A_0$ such that if $A \geqslant A_0$, then for $\varepsilon > 0$ sufficiently small, $\mu_{r,\alpha}(u) \leqslant \varepsilon$ and $\mu_{r,\alpha}(f) \leqslant \varepsilon$ imply $\mu_{r,\alpha}(u_n) \leqslant \varepsilon$.*

Assuming this lemma, one can prove Theorem 2 in exactly the same way as Theorem 2 in [2] was proved there. Since there is no change in this proof, we will say nothing further about it.

*Proof of the Lemma.* Exactly as in the proof of the lemma in §3 of [2], we have that if $\varepsilon > 0$ is sufficiently small, and $\mu_{r,\alpha}(f) < \varepsilon$, $\mu_{r,\alpha}(u) < \varepsilon$, then $\mu_{r,\alpha}(fu) < 3\varepsilon$.

From the definition of $u_0$, we have $\mu_{r,\alpha}(u_0) \leqslant A^r \mu_{r,\alpha}(fu)$ and $\mu_{r,\alpha}(u_0) \leqslant A^{r-1+\beta} \mu_{r,\alpha}(fu)$, if $\alpha$ is defined on all of $[0, \infty)$ and $\alpha(tx) \leqslant t^\beta \alpha(x)$, for $t \geqslant 1$.

From condition (6) on the mappings $P_{i,A}$, and the definition of $u_1, \ldots, u_n$, it follows that if $\varepsilon > 0$ is sufficiently small, then $\mu_{r,\alpha}(u_n) \leqslant 3C^n A^{r-n} \varepsilon$ and $\mu_{r,\alpha}(u_n) \leqslant 3C^n A^{r-n-1+\beta} \varepsilon$, if $\alpha$ is defined on all of $[0, \infty)$ and $\alpha(tx) \leqslant t^\beta \alpha(x)$, for $t \geqslant 1$.

Under the hypotheses of the lemma, the exponent of $A$ is negative, so by taking $A$ sufficiently large we may arrange that $3C^n A^{r-n} < 1$ or $3C^n A^{r-n-1+\beta} < 1$, according to the case. In either case, we have $\mu_{r,\alpha}(u_n) \leqslant \varepsilon$.                    Q.E.D.

## §3. Construction of the Mappings $P_{i,A}$

We consider the problem: given $u$ with support in $\mathrm{int}\, D_{i,A}$, find $v$ with support in $\mathrm{int}\, D_{i-1,A}$ such that $\tau_i u$ and $\tau_i v$ are conjugate. We also wish $v$ to satisfy an estimate of the form

$$\mu_{r,\alpha}(v) \leqslant CA^{-1} \mu_{r,\alpha}(u) \tag{1}$$

where $C$ is a constant (independent of $A$). This estimate should be satisfied for $u$ in a neighborhood of the identity. (This neighborhood may depend on $A$.)

Our method is similar to the method of [2, §5], where we solved the "inverse" problem. In particular, since $\mathrm{supp}\, u \subset \mathrm{int}\, D_{i,A} \subset \mathrm{int}\, D_{i-1,A}$, we may and do construct $h$ in the same way as there, provided $u$ is sufficiently close to the identity. Let $\bar{h}$ denote the unique diffeomorphism of $\mathbf{R}^n$ such that $\pi \bar{h} = h\pi$ and $\bar{h}$ is the identity on $\{x_i = 0\}$, where $\pi$ denotes the projection of $\mathbf{R}^n$ on $C_i$.

We let $B$ be a positive integer, which will be specified below. If $h$ is sufficiently $C^1$ close to the identity, $\tilde{h}$ will be close enough to the identity for us to define diffeomorphisms $\tilde{h}_1, ..., \tilde{h}_B$ of $\mathbf{R}^n$ by the formulae

$$\tilde{h}_1(x) - x = \frac{1}{B}(\tilde{h}(x) - x)$$

$$\tilde{h}_2(x) - x = \frac{1}{B-1}(\tilde{h} \circ \tilde{h}_1^{-1}(x) - x)$$

$$\tilde{h}_3(x) - x = \frac{1}{B-2}(\tilde{h} \circ \tilde{h}_1^{-1} \circ \tilde{h}_2^{-1}(x) - x)$$

$$...$$

$$\tilde{h}_{B-1}(x) - x = \frac{1}{2}(\tilde{h} \circ \tilde{h}_1^{-1} \circ \cdots \circ \tilde{h}_{B-2}^{-1}(x) - x)$$

$$\tilde{h}_B(x) - x = \tilde{h} \circ \tilde{h}_1^{-1} \circ \cdots \circ \tilde{h}_{B-1}^{-1}(x) - x .$$

Then $\tilde{h} = \tilde{h}_B \circ \tilde{h}_{B-1} \circ \cdots \circ \tilde{h}_1$.

Let $\zeta$ be a real $C^\infty$ function on $\mathbf{R}$, periodic of period 1, equal to 1 in a neighborhood of the integers, equal to zero in a neighborhood of the half-integers, and satisfying $0 \leqslant \zeta \leqslant 1$ everywhere.

For $j = 1, ..., B$, and $k = 0, 1$, we define $\tilde{h}_{jk}$ (for $k = 0, 1$) by the formulae $\tilde{h}_{j0}(x) - x = \zeta(x_i)(\tilde{h}_j(x) - x)$ where $x = (x_1, ..., x_n) \in \mathbf{R}^n$, and $\tilde{h}_{j1} = \tilde{h}_j \tilde{h}_{j0}^{-1}$. These will be diffeomorphisms if $h$ is $C^1$ close to id. We have $\tilde{h}_j = \tilde{h}_{j1} \tilde{h}_{j0}$ and $\tilde{h} = \tilde{h}_{B1} \tilde{h}_{B0} \tilde{h}_{B-1,1} \tilde{h}_{B-1,0} \cdots \tilde{h}_{1,1} \tilde{h}_{1,0}$.

This is the decomposition of $\tilde{h}$ we need in order to construct $v$.

CONSTRUCTION of $v$. We first construct two sequences $E_-^1, ..., E_-^B, E_+^1, ..., E_+^B$ of strips in $\mathbf{R}^n$. These are defined as follows. Let $a$ be the least half integer $\geqslant -2A$. Let

$$E_-^j = \{x \in \mathbf{R}^n : a + 3j - 3 \leqslant x_i \leqslant a + 3j - 2\}$$

$$E_+^j = \{x \in \mathbf{R}^n : a + 3j - 3/2 \leqslant x_i \leqslant a + 3j - 1/2\} .$$

We let $B$ be the greatest integer such that $a + 3B - 1/2 \leqslant 2A$.

In terms of increasing values of $x_i$, the strips occur in the following order: $E_-^1, E_+^1, E_-^2, E_+^2, ..., E_-^B, E_+^B$. Moreover they are disjoint, they all lie in the set $\{-2A \leqslant x_i \leqslant 2A\}$, and their sides are defined either by $x_i =$ half-integer (the $E_-^j$) or $x_i =$ integer (the $E_+^j$). They are squeezed as closely together as possible, compatibly with these properties.

We let $v \mid E_-^j$ be the unique diffeomorphism of $E_-^j$ onto itself such that $v \mid E_-^j = \tilde{h}_{j0} \mid E_-^j$. We let $v \mid E_+^j$ be the unique diffeomorphism of $E_+^j$ onto itself such that $v \mid E_+^j = \tilde{h}_{j1} \mid E_+^j$. We let

$$v \mid \left(\mathbf{R}^n - \bigcup_j (E_-^j \cap E_+^j)\right) = \text{id} .$$

If $u$ is sufficiently close to the identity this is a well defined diffeomorphism of $\mathbf{R}^n$.

Moreover, it is easily seen that $\Gamma_v^i = h$. Hence $\Gamma_v^i (\Gamma_u^i)^{-1} \in \mathscr{G}$, and it follows from the lemma in [2, §4] that $\tau_i u$ and $\tau_i v$ are conjugate in $\mathrm{Diff}(\mathbf{R}^n, r, \alpha)$, if $u$ is in $\mathrm{Diff}(\mathbf{R}^n, r, \alpha)$ and sufficiently $C^1$ close to the identity.

This completes the construction of $v$. Our estimate (1) will be proved in the next section.

We set $P_{i,A}(u) = v$. Of the properties (1)–(6) of $P_{i,A}$ listed in §2, properties (1)–(3) are obvious, we may arrange for (4) to hold by replacing the domain of $P_{i,A}$ by a smaller neighborhood of id, (5) is a consequence of the fact that $\tau_i u$ and $\tau_i v$ are conjugate, and (6) is equivalent to the inequality (1) of this section, which will be proved in the next section.

## §4. Estimate for $v$

In this section, we will complete the proof of Theorem 2 by proving the estimate (1) in §3. The proof is based on the following five estimates.

(1) If $u$ is $C^{r,\alpha}$ and sufficiently near the identity, then

$$\mu_{r,\alpha}(\Gamma_u) \leqslant 8 \mu_{r,\alpha}(u).$$

(2) If $u$ is $C^{r,\alpha}$ and sufficiently near the identity, then

$$\mu_{r,\alpha}(h) \leqslant 3 \mu_{r,\alpha}(\Gamma_u).$$

(3) There exists a constant $C_1 > 0$, independent of $A$, such that if $u$ is $C^{r,\alpha}$ and sufficiently near the identity, then

$$\mu_{r,\alpha}(\tilde{h}_i) \leqslant C_1 A^{-1} \mu_{r,\alpha}(h), \quad i = 1, \dots, B.$$

(4) There exists a constant $C_2$, independent of $A$, such that if $u$ is $C^{r,\alpha}$ and sufficiently close to the identity, then

$$\mu_{r,\alpha}(\tilde{h}_{i,j}) \leqslant C_2 \mu_{r,\alpha}(\tilde{h}_i), \quad j = 0, 1$$

(5) We have

$$\mu_{r,\alpha}(v) = \max \{\mu_{r,\alpha}(\tilde{h}_{ij})\}.$$

All but one of these estimates is obvious or is in [2, §6] in slightly different guise. Thus, estimate (1) is essentially a special case of (1) in [2, §6]. Here, supp $u \subset$ int $D_{i-1,A}$, whereas there, we had only the weaker condition supp $u \subset$ int $D_{i,A}$. This explains why we may omit $A$ from the right side of the inequality here: the width of $D_{i-1,A}$ in the $i$th coordinate is 4, while the width of $D_{i,A}$ is $4A$. The proof of (1) here is exactly the same as the proof of (1) in [2, §6].

Estimate (2) is exactly the same as (2) in [2, §6].

Estimate (3) is the new result. It will be proved below.

Estimate (4) is proved by the same argument that was used in Step 3 in [2, §6]. The equation (5) is obvious from the definitions.

From the estimates (1)–(5) of this section, we get that (1) of the previous section holds, with $C = 24 C_1 C_2$.

The estimate (3) is a consequence of the following lemma. Let $\mathscr{A}$ denote the group of $C^1$ diffeomorphisms $U$ of $\mathbf{R}^n$ such that $U(x) = x$ if $|x_j| \geqslant 2A$ for some $j \neq i$ or $x_i$ is an integer. For example, $\tilde{h}$ and the $\tilde{h}_i$ are in $\mathscr{A}$.

LEMMA. *Let* $0 < \lambda < 1$. *For any* $U \in \mathscr{A}$, *define* $V$ *by*

$$V(x) - x = \lambda (U(x) - x)$$

*and let* $W = UV^{-1}$ *(provided* $V^{-1}$ *exists, which is the case when* $U$ *is sufficiently close to the identity). Then there exists* $\delta > 0$ *(small) and* $C > 0$ *(large) such that for any* $U \in \mathscr{A}$ *satisfying* $\mu_{r,\alpha}(U) < \delta$, *we have*

$$\mu_{r,\alpha}(W) \leqslant (1 - \lambda)\, \mu_{r,\alpha}(U) + C \mu_{r,\alpha}(U)^2 . \tag{$*$}$$

We will prove this lemma below. First, however, we prove (3), assuming the lemma. Clearly

$$\mu_{r,\alpha}(\tilde{h}_1) = \frac{1}{B}\, \mu_{r,\alpha}(\tilde{h}) = \frac{1}{B}\, \mu_{r,\alpha}(h) .$$

Applying the lemma with $U = \tilde{h}$ and $\lambda = 1/B$, we get

$$\mu_{r,\alpha}(\tilde{h} \circ \tilde{h}_1^{-1}) \leqslant \frac{B-1}{B}\, \mu_{r,\alpha}(h) + O(\mu_{r,\alpha}(h)^2) .$$

Then, it is clear that

$$\mu_{r,\alpha}(\tilde{h}_2) \leqslant \frac{1}{B}\, \mu_{r,\alpha}(h) + O(\mu_{r,\alpha}(h)^2) .$$

Applying the lemma a second time, with $U = \tilde{h} \circ \tilde{h}_1^{-1}$ and $\lambda = (B-1)^{-1}$, we get

$$\mu_{r,\alpha}(\tilde{h} \circ \tilde{h}_1^{-1} \circ \tilde{h}_2^{-1}) \leqslant \frac{B-2}{B-1}\, \mu_{r,\alpha}(\tilde{h} \circ \tilde{h}_1^{-1}) + O(\mu_{r,\alpha}(\tilde{h} \circ \tilde{h}_1^{-1})^2)$$

$$\leqslant \frac{B-2}{B}\, \mu_{r,\alpha}(h) + O(\mu_{r,\alpha}(h)^2) .$$

Then, it is clear that $\mu_{r,\alpha}(\tilde{h}_3) \leqslant (1/B)\, \mu_{r,\alpha}(h) + O(\mu_{r,\alpha}(h)^2)$.

Continuing in this way, we see that we have

$$\mu_{r,\alpha}(\tilde{h}_i) \leqslant \frac{1}{B}\mu_{r,\alpha}(h) + O\left(\mu_{r,\alpha}(h)^2\right)$$

for $i = 1, 2, ..., B$. However, it is clear from the definition of $B$ that $B \geqslant \frac{1}{2}A$, so we get the estimate (3).

*Proof of the Lemma.* We give different proofs depending on whether $r = 1$ or $r > 1$. If $f$ is a mapping of $\mathbf{R}^n$ into itself, we define

$$v_r(f) = \sup_{x \in \mathbf{R}^n} \|D^r f(x)\|.$$

*Case $r > 1$.* We write $R_{f,g}$ for the sum of the "other terms" in formula (2) in §2 of [2]. Thus,

$$D^r(f \circ g) = (D^r f \circ g) \cdot (Dg)^r + (Df \circ g) \cdot D^r g + R_{f,g}. \tag{6}$$

If $f$ and $g$ are $C^r$, then $R_{f,g}$ is $C^1$ and there exists $\delta > 0$ and $C > 0$ such that if $f, g \in \mathscr{A}$ and $v_r(f), v_r(g) < \delta$, then

$$v_1(R_{f,g}) \leqslant Cv_r(f)\, v_r(g).$$

Applying (6) to $g = f^{-1}$, we get

$$D^r(f^{-1}) = -\left((Df)^{-1} \circ f^{-1}\right) \cdot (D^r f \circ f^{-1}) \cdot \left((Df)^{-1} \circ f^{-1}\right)^r + R_f, \tag{7}$$

where there exists $\delta > 0$ and $C > 0$ such that if $f \in \mathscr{A}$ and $v_r(f) < \delta$, then

$$v_1(R_f) \leqslant Cv_r(f)^2.$$

Then

$$D^r W = (D^r U \circ V^{-1}) \cdot (DV^{-1})^r + (DU \circ V^{-1}) \cdot D^r V^{-1} + R_{U,1}$$
$$= (D^r U \circ V^{-1}) - \lambda(D^r U \circ V^{-1}) + S_U + R_{U,2}$$

where

$$S_U = (D^r U \circ V^{-1}) \cdot (DV^{-1})^r - D^r U \circ V^{-1}$$
$$+ (D^r V \circ V^{-1} - ((DU \circ V^{-1}) \cdot ((DV)^{-1} \circ V^{-1}) \cdot (D^r V \circ V^{-1}) \cdot ((DV)^{-1} \circ V^{-1})^r))$$

and $R_{U,i}$ $(i = 1, 2)$ has the property that there exists $\delta > 0$ and $C > 0$ such that if $U \in \mathscr{A}$ and $v_r(U) < \delta$, then $v_1(R_{U,i}) < Cv_r(U)^2$.

Then it is easily seen that there exist $\delta > 0$ and $C > 0$, such that if $U \in \mathscr{A}$ and $\mu_{r,\alpha}(U) < \delta$, we have

$$\mu_\alpha(S_U + R_{U,2}) \leqslant C\mu_{r,\alpha}(U)^2.$$

The lemma (in the case $r > 1$) follows immediately.

*Case $r = 1$.* We have

$$DW = (DU \circ V^{-1}) \cdot D(V^{-1}) = (DU \circ V^{-1}) \cdot ((DV)^{-1} \circ V^{-1}) = (DU \cdot (DV)^{-1}) \circ V^{-1}.$$

(8)

Then

$$(DV)^{-1} = (I + (DV - I))^{-1} = I - (DV - I) + (DV - I)^2 - \cdots$$
$$= I - \lambda(DU - I) + \lambda^2(DU - I)^2 - \cdots .$$

Therefore

$$DU \cdot (DV)^{-1} = I + (1 - \lambda)(DU - I) + a_2(DU - I)^2 + a_3(DU - I)^3 + \cdots$$

(9)

where $a_2 z^2 + a_3 z^3 + \cdots$ is a convergent power series (for $|z| < \lambda^{-1}$). The lemma, in the case $r = 1$, follows immediately from formulas (8) and (9).

## §5. Application to Haefliger's Classifying Space

This is just like §7 of [2]. We get that $F\Gamma_n^r$ is $(n + 1)$-connected if $n \geqslant r \geqslant 1$. Likewise $F\Gamma_n^{r; \alpha}$ is $(n + 1)$-connected under the hypotheses of Theorem 2. These assertions follow in exactly the same way from the main results of this paper as the assertions in §7 of [2] followed from the main results of [2], so we will not repeat the proofs here.

## REFERENCES

[1] EPSTEIN, D. B. A., *The Simplicity of Certain Groups of Homeomorphisms*, Compositio Math. *22* (1970), 165–173.

[2] MATHER, J. N., *Commutators of Diffeomorphisms*, to appear in Comment Math. Helv.

*Dept. of Mathematics*
*Harvard University*
*1 Oxford Street*
*Cambridge, Mass. 02138*
*U.S.A.*