Hadamard and Perron

JWR

October 18, 1999

On page 23 of his famous monograph [2], D. V. Anosov writes

> Every five years or so, if not more often, someone 'discovers' the theorem of Hadamard and Perron proving it either by Hadamard's method or Perron's. I myself have been guilty of this.

1 Notation

If (X, d_X) and (Y, d_Y) are metric spaces and $T : X \to Y$ is a map then the **Lipschitz constant** of T is the quantity

$$\text{lip}(T) = \sup \left\{ \frac{d_Y(T(x_1), T(x_2))}{d_X(x_1, x_2)} : x_1, x_2 \in X \; x_1 \neq x_2 \right\}.$$

If X and Y are Banach spaces and $q : X \to Y$ satisfies $q(0) = 0$ then the **Lipschitz constant at** 0 is the quantity

$$\|q\|_* = \sup \left\{ \frac{\|q(x)\|_Y}{\|x\|_X} : x \in X \setminus 0 \right\}$$

2 Lipschitz Inverse Function Theorem

Theorem 2.1. *Suppose that X is a Banach space and*

$$f : X \to X$$

has form

$$f(x) = x + r(x)$$

where

$$\text{lip}(r) < 1.$$

Then f is a lipeomorphism and

$$\text{lip}(f^{-1}) \leq (1 - \text{lip}(r))^{-1}.$$

Proof. Choose $u \in X$ and define $\xi_u : X \to X$ by

$$\xi_u(x) = u - r(x).$$

Thus $f(x) = u$ if and only if $\xi_u(x) = x$. As $\mathrm{lip}(\xi_u) = \mathrm{lip}(r) < 1$ there is a unique fixed point x (for each choice of u) by the Contraction Mapping Principle; thus f is bijective. To estimate the Lipschitz constant of f^{-1} we compute

$$\begin{aligned}
\|f^{-1}(u) - f^{-1}(v)\| &= \|x - y\| \\
&= \|\xi_u(x) - \xi_v(y)\| \\
&= \|(u - r(x)) - (v - r(y))\| \\
&\leq \|u - v\| + \mathrm{lip}(r)\|x - y\| \\
&= \|u - v\| + \mathrm{lip}(r)\|f^{-1}(u) - f^{-1}(v)\|
\end{aligned}$$

where $x = f^{-1}(u)$ and $y = f^{-1}(v)$. This proves

$$\|f^{-1}(u) - f^{-1}(v)\| \leq (1 - \mathrm{lip}(r))^{-1}\|u - v\|$$

as required. □

3 The Resolvent Inequality

Lemma 3.1. *Now suppose $f_i : X \to X$ for $i = 1, 2$ both have form*

$$f_i(x) = x + r_i(x)$$

where $\mathrm{lip}(r_i) < 1$ and now assume also that

$$r_i(0) = 0.$$

Then

$$\|f_1^{-1} - f_2^{-1}\|_* \leq c_1 c_2 \|r_1 - r_2\|_*$$

where

$$c_i = (1 - \mathrm{lip}(r_i))^{-1}.$$

Proof. For $u \in X$ we have

$$f_i^{-1}(u) = u - r_i(f_i^{-1}(u))$$

so that

$$\|f_1^{-1}(u) - f_2^{-1}(u)\| \leq \tau + \sigma \tag{!}$$

where

$$\tau = \|r_1(f_1^{-1}(u) - r_1(f_2^{-1}(u)\|$$

and

$$\sigma = \|(r_1 - r_2)(f_2^{-1}(u))\|.$$

for τ we have

$$\tau \leq \text{lip}(r_1)\|f_1^{-1}(u) - f_2^{-1}(u)\|$$

so that (!) becomes

$$(1 - \text{lip}(r_1))\|f_1^{-1}(u) - f_2^{-1}(u)\| \leq \sigma.$$

For σ we have

$$\sigma \leq \|r_1 - r_2\|_* \, \|f_2^{-1}(u)\|.$$

But $\text{lip}(f_2^{-1}) \leq c_2$ by the Lipschitz Inverse Function Theorem 2.1 so

$$\|f_2^{-1}(u)\| \leq c_2\|u\|$$

as $f_2^{-1}(0) = 0$. Combining these inequalities gives the estimate

$$\|f_1^{-1}(u) - f_2^{-1}(u)\| \leq c_1 c_2 \|r_1 - r_2\|_* \, \|u\|$$

as required. □

For linear transformations the Lipschitz constant and the Lipschitz constant from 0 are both equal to the operator norm and the inequality just proved takes the form

$$\|T_1^{-1} - T_2^{-1}\| \leq \|T_1^{-1}\| \, \|T_2^{-1}\| \, \|T_1 - T_2\|.$$

This in turn follows immediately from the identity

$$T_1^{-1} - T_2^{-1} = T_1^{-1}(T_2 - T_1)T_2^{-1}.$$

When T_1 and T_2 commute this is

$$T_1^{-1} - T_2^{-1} = (T_2 - T_1)T_1^{-1}T_2^{-1}$$

and a special case of this is the Resolvent Identity

$$R_\lambda - R_\mu = (\mu - \lambda)R_\lambda R_\mu$$

(where $T_1 = \lambda - T$, $T_2 = \mu - T$, $R_\lambda = T_1^{-1}$, and $R_\mu = T_2^{-1}$.)

4 Graph Transform Lemma

The arguments in this section come from [4]. Let X, Y, U, V be Banach spaces (ultimately we take $X = U$ and $Y = V$) and

$$F : X \times Y \to U \times V$$

have the form

$$F(x, y) = (A(x) + R(x, y), H(x, y))$$

where $A : X \to U$, $R : X \times Y \to U$, $H : X \times Y \to V$.

Lemma 4.1. *Assume that the map $A : X \to U$ is bijective and*

$$\text{lip}(A^{-1})\text{lip}(R) < 1$$

Then for each $\alpha : X \to Y$ satisfying

$$\text{lip}(\alpha) \leq 1$$

there exists a (necessarily unique because the graph of a function determines the function uniquely) $\beta : U \to V$ satisfying

$$\text{graph}(\beta) = F(\text{graph}(\alpha))$$

The map β is called the **graph transform** *of α by F.*

Proof. We can rewrite this as follows. Define

$$\phi : X \to U$$

depending on α by

$$\phi = A + R \circ (1, \alpha).$$

Then we must solve

$$\beta \circ \phi = H \circ (1, \alpha).$$

This suggests inverting ϕ. Now ϕ has the form

$$\phi = (1 + R \circ (1, \alpha) \circ A^{-1}) \circ A.$$

where A is bijective. The term $R \circ (1, \alpha) \circ A^{-1}$ has a Lipschitz constant satisfying

$$\text{lip}(R \circ (1, \alpha) \circ A^{-1}) \leq \text{lip}(R)\text{lip}(A^{-1}) < 1$$

since $\text{lip}((1, \alpha)) = 1$. (The norm on $U \times V$ is $\|(u, v)\| = \max(\|u\|, \|v\|)$.) According to the Lipschitz Inverse Function Theorem 2.1, the map ϕ is a bijection so β is given by

$$\beta = H \circ (1, \alpha) \circ \phi^{-1}.$$

□

5 Hadamard

The arguments in this section come from [4]. Let X and Y be Banach spaces and

$$F : X \times Y \to X \times Y$$

have the form

$$F(x, y) = (A(x) + R(x, y), H(x, y))$$

where

$$A : X \to X, \qquad R : X \times Y \to X, \qquad H : X \times Y \to Y.$$

Theorem 5.1. *Assume that the map $A : X \to X$ is bijective, that*

$$\mathrm{lip}(A^{-1})\,\mathrm{lip}(H) < 1 \tag{1}$$

and that

$$A(0) = 0, \quad R(0,0) = 0, \quad H(0,0) = 0. \tag{2}$$

Assume also

$$\mathrm{lip}(R) < \epsilon \tag{3}$$

where $\epsilon > 0$ is a small constant depending only on $\mathrm{lip}(H)$ and $\mathrm{lip}(A^{-1})$ defined explicitly in the proof. Then there exists a unique $\alpha : X \to Y$ satisfying

$$F(\mathrm{graph}(\alpha)) = \mathrm{graph}(\alpha)$$

and

$$\alpha(0) = 0, \quad \mathrm{lip}(\alpha) \leq 1.$$

Proof. According to the Graph Transform Lemma 4.1 each $\alpha : X \to Y$ with $\mathrm{lip}(\alpha) \leq 1$ determines a unique $\beta : X \to Y$ satisfying $F(\mathrm{graph}(\alpha)) = \mathrm{graph}(\beta)$. This *graph transform* is defined by

$$\beta := H \circ (1, \alpha) \circ \phi^{-1}, \qquad \phi := A + R \circ (1, \alpha). \tag{4}$$

To satisfy the hypothesis of Lemma 4.1 we require only $\epsilon < \mathrm{lip}(A^{-1})^{-1}$. Denote by $\mathcal{G}$ the space of all $\alpha : X \to Y$ such that $\alpha(0) = 0$ and $\mathrm{lip}(\alpha) \leq 1$. We first show that the graph transform maps $\mathcal{G}$ to itself. By hypothesis (2) we have $F(0,0) = (0,0)$ so $\beta(0) = 0$ follows from $\alpha(0) = 0$. Equation (4) gives the estimate

$$\mathrm{lip}(\beta) \leq \mathrm{lip}(H)\mathrm{lip}(\phi^{-1}).$$

But $\phi = (1 + R \circ (1, \alpha) \circ A^{-1}) \circ A$ and hence by the Lipschitz Inverse Function Theorem 2.1 we have the estimate

$$\mathrm{lip}(\phi^{-1}) \leq \mathrm{lip}(A^{-1})C$$

where

$$C = (1 - \mathrm{lip}(R)\mathrm{lip}(A^{-1}))^{-1}. \tag{5}$$

Thus

$$\mathrm{lip}(\beta) \leq \mathrm{lip}(H)\mathrm{lip}(A^{-1})C.$$

Now C is near 1 by (3) so that $\mathrm{lip}(\beta) \leq 1$ follows from the hypothesis (1).

Next we show that the map which assigns to $\alpha \in \mathcal{G}$ its graph transform $\beta \in \mathcal{G}$ is a strict contraction in the norm $\|\cdot\|_*$, i.e. there is a constant $\lambda < 1$ such that

$$\|\beta_1 - \beta_2\|_* \leq \lambda \|\alpha_1 - \alpha_2\|_* \tag{6}$$

whenever $\alpha_i(0) = 0$ and $\mathrm{lip}(\alpha_1) \leq 1$ (for $i = 1, 2$) and β_i is the graph transform of α_i. Abbreviate

$$\phi_i := A + R \circ (1, \alpha_i), \qquad r_i = R \circ (1, \alpha_i) \circ A^{-1}, \qquad f_i = 1 + r_i,$$

so $\phi_i = f_i \circ A$ and hence $\phi_i^{-1} = A^{-1} \circ f_i^{-1}$. For $u \in X$ we have

$$\|\beta_1(u) - \beta_2(u)\| \leq \sigma + \tau$$

where

$$\sigma := \|H \circ (1, \alpha_1) \circ \phi_1^{-1}(u) - H \circ (1, \alpha_1) \circ \phi_2^{-1}(u)\|$$

and

$$\tau := \|H \circ (1, \alpha_1) \circ \phi_2^{-1}(u) - H \circ (1, \alpha_2) \circ \phi_2^{-1}(u)\|.$$

Recall that $\mathrm{lip}((1, \alpha)) = 1$ and $C = (1 - \mathrm{lip}(R)\mathrm{lip}(A^{-1}))^{-1}$. We estimate σ:

$$\begin{aligned}
\sigma &\leq \mathrm{lip}(H)\|\phi_1^{-1}(u) - \phi_2^{-1}(u)\| \\
&\leq \mathrm{lip}(H)\mathrm{lip}(A^{-1})\|f_1^{-1}(u) - f_2^{-1}(u)\| \\
&\leq \mathrm{lip}(H)\mathrm{lip}(A^{-1})C^2\|r_1(u) - r_2(u)\| \\
&\leq \mathrm{lip}(H)\mathrm{lip}(A^{-1})C^2\mathrm{lip}(R)\|\alpha_1(A^{-1}(u)) - \alpha_2(A^{-1}(u))\| \\
&\leq \mathrm{lip}(H)\mathrm{lip}(A^{-1})C^2\mathrm{lip}(R)\|\alpha_1 - \alpha_2\|_* \|A^{-1}(u)\| \\
&\leq \mathrm{lip}(H)\mathrm{lip}(A^{-1})^2C^2\mathrm{lip}(R)\|\alpha_1 - \alpha_2\|_* \|u\|
\end{aligned}$$

We estimate τ:

$$\begin{aligned}
\tau &\leq \mathrm{lip}(H)\|(\alpha_1 - \alpha_2)(\phi_2^{-1}(u))\| \\
&\leq \mathrm{lip}(H)\|\alpha_1 - \alpha_2\|_* \|\phi_2^{-1}(u))\| \\
&\leq \mathrm{lip}(H)\|\alpha_1 - \alpha_2\|_* \mathrm{lip}(A^{-1})C\|u\|
\end{aligned}$$

Combining the estimates for σ and τ gives

$$\|\beta_1(u) - \beta_2(u)\| \leq \lambda\|\alpha_1 - \alpha_2\|_* \|u\|$$

where $\lambda < 1$ for $\mathrm{lip}(R)$ small enough. This proves (6).

The space $\mathcal{G}$ is a complete metric space in the metric $d(\alpha_1, \alpha_2) = \|\alpha_1 - \alpha_2\|_*$. This is because convergence in the norm $\|\cdot\|_*$ implies pointwise convergence and uniform Lipschitz estimates are preserved under pointwise limits. Hence the theorem follows from the Contraction Mapping Principle. □

6 Perron

The arguments in this section come from [8]. Let X and Y be Banach spaces and

$$F : X \times Y \to X \times Y$$

have the form

$$F(x, y) = (K(x, y), By + S(x, y))$$

where

$$K : X \times Y \to X, \qquad B : Y \to Y, \qquad S : X \times Y \to Y.$$

Theorem 6.1. *Assume that the map $B : Y \to Y$ is linear and invertible, that*

$$\|B^{-1}\|\Big(\mathrm{lip}(K) + \mathrm{lip}(S)\Big) < 1, \tag{7}$$

and that

$$K(0,0) = 0, \quad S(0,0) = 0.$$

Choose $\delta \in (0,1)$ so small that

$$\lambda := \|B^{-1}\|\Big((1+\delta)\mathrm{lip}(K) + \mathrm{lip}(S)\Big) < 1 \tag{8}$$

and assume that $\mathrm{lip}(S)$ is so small that

$$\|B^{-1}\|\,(\delta\,\mathrm{lip}(K) + \mathrm{lip}(S)) < \delta. \tag{9}$$

Then there exists a unique $\alpha : X \to Y$ satisfying

$$F(\mathrm{graph}(\alpha)) = \mathrm{graph}(\alpha)$$

and

$$\alpha(0) = 0, \quad \mathrm{lip}(\alpha) \le \delta.$$

Proof. The condition $F(\mathrm{graph}(\alpha)) = \mathrm{graph}(\alpha)$ may be written in the form

$$B\alpha + S \circ (1,\alpha) = \alpha \circ K \circ (1,\alpha)$$

which is in turn equivalent to

$$\alpha = \Gamma(\alpha)$$

where $\Gamma(\alpha) : X \to Y$ is defined by

$$\Gamma(\alpha) = B^{-1}\Big(\alpha \circ K \circ (1,\alpha) - S \circ (1,\alpha)\Big). \tag{10}$$

Lemma 6.2. *Suppose $\alpha : X \to Y$ satisfies $\alpha(0) = 0$ and $\mathrm{lip}(\alpha) \le \delta$ Then $\beta = \Gamma(\alpha)$ satisfies these same conditions: $\beta(0) = 0$ and $\mathrm{lip}(\beta) \le \delta$.*

Proof. Since $K(0,0) = 0$, $S(0,0) = 0$, B^{-1} is linear, and $\alpha(0) = 0$ it follows that $\beta(0) = 0$ from the definition of β. For the Lipschitz estimate we compute

$$\begin{aligned}
\mathrm{lip}(\beta) &\le \|B^{-1}\|\Big(\mathrm{lip}(\alpha) \circ K \circ (1,\alpha) - S \circ (1,\alpha)\Big) \\
&\le \|B^{-1}\|\Big(\mathrm{lip}(\alpha \circ K \circ (1,\alpha)) + \mathrm{lip}(S \circ (1,\alpha))\Big) \\
&\le \|B^{-1}\|\Big(\mathrm{lip}(\alpha)\mathrm{lip}(K)\mathrm{lip}((1,\alpha)) + \mathrm{lip}(S)\mathrm{lip}((1,\alpha))\Big) \\
&= \|B^{-1}\|\Big(\mathrm{lip}(\alpha)\mathrm{lip}(K) + \mathrm{lip}(S)\Big) \\
&\le \|B^{-1}\|\Big(\delta\,\mathrm{lip}(K) + \mathrm{lip}(S)\Big) \\
&\le \delta.
\end{aligned}$$

Here we have use the fact that $\text{lip}(1, \alpha) = 1$ (from $\text{lip}(\alpha) \leq \delta \leq 1$). This completes the proof of lemma 6.2. □

Lemma 6.3. *The map Γ is a contraction in the $\|\cdot\|_*$ norm. More precisely, if $\alpha_i : X \to Y$ for $i = 1, 2$ satisfies $\alpha_i(0) = 0$ and $\text{lip}(\alpha_i) \leq \delta$ and if $\beta_i = \Gamma(\alpha_i)$, then*

$$\|\beta_1 - \beta_2\|_* \leq \lambda \|\alpha_1 - \alpha_2\|_*.$$

Proof. Choose $x \in X$. Then

$$\|\beta_1(x) - \beta_2(x)\| \leq \|B^{-1}\|(\sigma + \tau + \rho)$$

where

$$\begin{aligned}
\sigma &= \|\alpha_1(K(x, \alpha_1(x))) - \alpha_1(K(x, \alpha_2(x)))\| \\
\tau &= \|\alpha_1(K(x, \alpha_2(x))) - \alpha_2(K(x, \alpha_2(x)))\| \\
\rho &= \|S(x, \alpha_1(x)) - S(x, \alpha_2(x))\|
\end{aligned}$$

For σ we have

$$\begin{aligned}
\sigma &\leq \text{lip}(\alpha_1)\text{lip}(K)\|\alpha_1(x) - \alpha_2(x)\| \\
&\leq \delta\text{lip}(K)\|\alpha_1(x) - \alpha_2(x)\|
\end{aligned}$$

For τ we have

$$\begin{aligned}
\tau &\leq \|\alpha_1 - \alpha_2\|_* \,\text{lip}(K) \max(\|x\|, \|\alpha_2(x)\|) \\
&= \|\alpha_1 - \alpha_2\|_* \,\text{lip}(K)\|x\|
\end{aligned}$$

since $\alpha_2(0) = 0$,$K(0, 0) = 0$, and $\text{lip}(\alpha_2) \leq \delta \leq 1$. For ρ we have

$$\rho \leq \text{lip}(S)\|\alpha_1(x) - \alpha_2(x)\|$$

Now divide by $\|x\|$ and take the supremum over x. This completes the proof of lemma 6.3. □

Now the space $\mathcal{G}_\delta$ of all $\alpha : X \to Y$ with $\alpha(0) = 0$ and $\text{lip}(\alpha) \leq \delta$ form a complete metric space in the metric $d(\alpha_1, \alpha_2) = \|\alpha_1 - \alpha_2\|_*$. This is because convergence in the norm $\|\cdot\|_*$ implies pointwise convergence and uniform Lipscitz estimates are preserved under pointwise limits. According to lemmas 6.2 and 6.3 the map Γ is a strict contraction on this space. This completes the proof of Theorem 6.1. □

7 Smoothness

Lemma 7.1 (Fiber Contraction Principle)**.** *Let $\mathcal{G}$ be a topological space, $\mathcal{H}$ be a complete metric space, and $\Phi : \mathcal{G} \times \mathcal{H} \to \mathcal{G} \times \mathcal{H}$ a map of form*

$$\Phi(g, h) = (\Gamma(g), \Delta_g(h)).$$

Assume

(i) *$\Gamma : \mathcal{G} \to \mathcal{G}$ has an attractive fixed point.*

(ii) *For each $h \in \mathcal{H}$ the map $\mathcal{G} \to \mathcal{H} : g \mapsto \Delta_g(h)$ is continuous.*

(iii) *There exists $\rho \in [0,1)$ such that $d(\Delta_g(h_1), \Delta_g(h_2)) \leq \rho d(h_1, h_2)$ for all $g \in \mathcal{G}$ and $h_1, h_1 \in \mathcal{H}$.*

Then $\Phi : \mathcal{H} \to \mathcal{H}$ has an attractive fixed point.

Proof. Let $\bar{g}$ be the fixed point of Γ and $\bar{h}$ be the fixed point of the contraction map $\Delta_{\bar{g}}$. Choose $(g,h) \in \mathcal{H} \times \mathcal{H}$ and let $(g_n, h_n) = \Phi^n(g,h)$, and $\sigma_n = d(h_n, \bar{h})$. Then $g_n \to \bar{g}$ as $n \to \infty$ by (i); we must show $\sigma_n \to 0$ as $n \to \infty$. Now by (iii)

$$\begin{aligned} d(h_{n+1}, \bar{h}) &= d(\Delta_{g_n}(h_n), \bar{h}) \\ &\leq d(\Delta_{g_n}(h_n), \Delta_{g_n}(\bar{h})) + d(\Delta_{g_n}(\bar{h}), \bar{h}) \\ &\leq \rho d(h_n, \bar{h}) + d(\Delta_{g_n}(\bar{h}), \bar{h}) \end{aligned}$$

i.e. $\sigma_{n+1} \leq \rho\sigma_n + \delta_n$ where $\delta_n = d(\Delta_{g_n}(\bar{h}), \bar{h})$. Hence by induction

$$\sigma_n \leq \rho^n \sigma_0 + \sum_{k=0}^{n-1} \rho^{n-1-k} \delta_k.$$

For $0 < m < n$ we have

$$\sigma_n \leq \rho^n \sigma_0 + C \sum_{k=0}^{m} \rho^{n-1-k} + c \sum_{k=m+1}^{n-1} \rho^{n-1-k} \leq \rho^n \sigma_0 + \frac{C\rho^{n-1-m} + c}{1 - \rho}$$

where $C = \max_{0 \leq k \leq m} \delta_k \leq \max_k \delta_k$ and $c = \max_{m < k \leq n} \delta_k \leq \max_{m < k} \delta_k$. But $\Delta_{\bar{g}}(\bar{h}) = \bar{h}$ so $\delta_n = d(\Delta_{g_n}(\bar{h}), \Delta_{\bar{g}}(\bar{h}))$ so $\delta_n \to 0$ as $n \to \infty$ by (i) and (ii). Hence σ_n is small if $n > 2m$ and m is large. □

Theorem 7.2. *Assume that $F : X \times Y \to X \times Y$ is a C^k diffeomorphism on a product of Banach spaces of form*

$$F(x,y) = \big(Ax + R(x,y), By + S(x,y)\big)$$

where

$$R(0,0) = 0, \quad S(0,0) = 0, dR(0,0) = 0, \quad dS(0,0) = 0$$

and $A \in L(X,X)$ and $B \in L(Y,Y)$ are invertible linear maps satisfying the spectral gap condition

$$\|A^{-1}\|^i \, \|B\| < 1, \qquad i = 1, 2, \ldots, k. \tag{11}$$

Then for each $k = 1, 2, \ldots$ there exists $\delta_k > 0$ such that if $\|R\|_k + \|S\|_k < \delta_k$, then the (hypotheses of Theorem 5.1 hold and) the map $\alpha : X \to Y$ defined in Theorem 5.1 is C^k.

Proof. First we show that with $H(x, y) = By + S(x, y)$ the hypotheses of Theorem 5.1 hold for $\|R\|_1 + \|S\|_1 < \delta_1$ and $\delta = \delta_1$ sufficiently small. This follows easily from $\mathrm{lip}(R) \leq \|R\|_1$, $\mathrm{lip}(S) \leq \|S\|_1$, $\mathrm{lip}(A^{-1}) = \|A^{-1}\|$, $\mathrm{lip}(B) = \|B\|$. and the fact that (11) (with $i = 1$) implies that (1) and (3) hold for $R = 0$ and $S = 0$.

Let $\mathcal{G}$ be the metric space of all maps $\alpha : X \to Y$ with $\alpha(0) = 0$ and $\mathrm{lip}(\alpha) \leq 1$ with metric $d(\alpha_1, \alpha_2) = \|\alpha_1 - \alpha_2\|_*$ as in the proof of Theorem 5.1. Define $\Gamma : \mathcal{G} \to \mathcal{G}$ by

$$\Gamma(\alpha) := B\alpha \circ \phi^{-1} + S \circ (1, \alpha) \circ \phi^{-1}, \qquad \phi := A + R \circ (1, \alpha),$$

i.e. $\Gamma(\alpha)$ is the graph transform of α as in (4) in the proof of Theorem 5.1. Assume that α is smooth. Then so is $\beta = \Gamma(\alpha)$. Differentiating $\beta(x)$ with respect to x gives

$$d\beta(x) = \Big(Bd\alpha(x') + dS(x', y')(1, d\alpha(x'))\Big) d\phi(x)^{-1}$$

where $x' = \phi^{-1}(x)$ and $y' = \alpha(\phi^{-1}(x))$. From $d\phi^{-1}(x) = d\phi(\phi^{-1}(x))^{-1}$ we get

$$d\phi^{-1}(x) = \Big(A + dR(x', y')(1, d\alpha(x')\Big)^{-1}$$

In both formulas the derivative of α is evaluated at x' so combining gives a formula of form

$$d\beta(x) = \mathcal{J}_1(x', \alpha(x'), d\alpha(x')).$$

Successive differentiation gives a formula of form

$$j^k\beta(x) = \mathcal{J}_k(x', \alpha(x'), j^k\alpha(x'))$$

where j^k indicates the k-jet, i.e.

$$j^k\alpha(x) = (d\alpha(x), d^2\alpha(x), \cdots, d^k\alpha(x)).$$

The Banach space of bounded linear maps from X to Y is denoted by $L(X, Y)$, the space $L^k(X, Y)$ bounded k-multilinear maps from X^k to Y is defined inductively by the equations

$$L^0(X, Y) = Y, \qquad L^{k+1}(X, Y) = L(X, L^k(X, Y))$$

and the subspace of symmetric maps is denoted by $L^k_{\mathrm{sym}}(X, Y) \subset L^k(X, Y)$. Define

$$P^k(X, Y) = L(X, Y) \times L^2_{\mathrm{sym}}(X, Y) \times \cdots \times L^k_{\mathrm{sym}}(X, Y)$$

so that the k-jet of a C^k map $\alpha : X \to Y$ is a map $j^k\alpha : X \to P^k(X, Y)$. Let $\mathcal{H}$ be the Banach space of all bounded continuous maps $h : X \to P^k(X, Y)$ with the sup norm. For $\alpha \in \mathcal{G}$ define $\Delta_\alpha : \mathcal{H} \to \mathcal{H}$ by

$$\Delta_\alpha(h)(x) = \mathcal{J}_k(x', \alpha(x'), h(x')).$$

If R and S vanish identically, we have that $\Gamma(\alpha)(x) = B\alpha(A^{-1}x)$ and hence $\Delta_\alpha(h)(x) = BA_*(h(A^{-1}x))$ where $A_* : P^k(X,Y) \to P^k(X,Y)$ is defined by

$$A_*(p_1, p_2, \ldots, p_k)(\hat{x}) = \big(p_1(A^{-1}\hat{x}), p_2(A^{-1}\hat{x})^2, \ldots, p_k(A^{-1}\hat{x})^k\big).$$

The spectral gap condition (11) guarantees that Δ_α is a contraction uniformly in α when R and S vanish identically and hence if $\|R\|_k$ znd $\|S\|_k$ are sufficiently small. By the Fiber Contraction Principle the map $\Phi(\alpha, h) = (\Gamma(\alpha), \Delta_\alpha(h))$ has an attractive fixed point. But if $(\alpha_n, h_n) = \Phi^n(\alpha, h)$, α is C^k. and $h = j^k\alpha$, then $h_n = j^k\alpha_n$, so the first k derivatives of α_n converge uniformly and the limit is also C^k. □

Remark 7.3. A similar proof is given in [8] using the map Γ from the proof of Theorem 6.1 in place of the graph transform.

Example 7.4. We show that Theorem 7.2 fails for $k = \infty$. Choose a smooth function $S : \mathbb{R} \to \mathbb{R}$ with $S(0) = dS(0) = 0$. Take $X = \mathbb{R}^2$, $Y = \mathbb{R}$, $\mu = 1/2$, and let $F : X \times Y \to X \times Y$ be a diffeomorphism such that

$$F(\ell, x, y) = (\ell, \ell x, \mu y + S(x))$$

in a neighborhood U of $(\ell, x, y) = (1, 0, 0)$. (It follows that this point is a fixed point of F.) We show that F has no invariant manifold graph(α) where $\alpha : X \to Y$ is smooth and $\alpha(1, 0) = 0$. The condition $F(\text{graph}(\alpha)) = \text{graph}(\alpha)$ implies that

$$\mu\alpha(\ell, x) + h(x) = \alpha(\ell, \ell x)$$

for $(\ell, x, \alpha(\ell, x)) \in U$. If $\alpha \in C^\infty$ this implies

$$\mu \frac{\partial^n \alpha}{\partial x^n}(\ell, x) + d^n S(x) = \ell^n \frac{\partial^n \alpha}{\partial x^n}(\ell, \ell x)$$

Taking $(\ell, x) = (\mu^{1/n}, 0)$ and n is so large that $(\ell, x, \alpha(\ell, x)) \in U$ shows that $d^n S(0) = 0$. Thus if $dS^n(0) \neq 0$ for infinitely many n no such smooth α exists. By multiplying the nonlinear terms by a smooth cuttoff function and rescaling we can construct, for any positive integer k, a smooth diffeomorphism

$$F(\ell, x, y) = (\ell, x, \mu y) + Q(\ell, x, y)$$

where Q vanishes to first order at the fixed point, the C^k norm of Q is as small as desired, and the invariant manifold $y = \alpha(\ell, x)$ of Theorem 7.2 is not smooth. (This example comes from Lanford's lecture notes [7] where it was adapted it from [9].)

Now we sketch the proof of smoothness from [5]. It is a discrete version of the proofs of [3] and [6], but unlike those, it explicitly uses the infinite dimensional implicit function theorem.

Let F be a C^r map $(r \geq 1)$ from a Banach space to itself with a hyperbolic fixed point at the origin. Thus after a suitable linear change of co-ordinates and choice of norms, the Banach space is a product $X \times Y$ of closed subspaces and $F\colon X \times Y \to X \times Y$ has the form

$$F(x, y) = (Ax + R(x, y), By + S(x, y))$$

where A is a linear contraction on X, B^{-1} is a linear contraction on Y, and R, S, DR, DS vanish at the origin. Give $X \times Y$ the product norm and let $B_{X\times Y} = B_X \times B_Y$ be a small ball about the origin. The stable manifold theorem asserts the existence of a (necessarily unique) C^r map $\alpha\colon B_X \to B_Y$ such that the graph of α is precisely the set of all $z \in B_{X\times Y}$ such that $F^n(z) \in B_{X\times Y}$ for all $n \geq 0$. It follows that $F^n(z) \to 0$ as $n \to \infty$ for $z \in \operatorname{graph}(\alpha)$.

Choose $z = (x, \alpha(x)) \in \operatorname{graph}(\alpha)$ and define $(x_n, y_n) = z_n = F^n(z)$ so that

$$x_{n+1} = Ax_n + R(z_n), \qquad y_{n+1} = By_n + S(z_n). \tag{$\dagger$}$$

It follows easily that

$$x_n = A^n x + \sum_{\nu=0}^{n-1} A^{n-1-\nu} R(z_\nu), \qquad y_n = -\sum_{\nu=n}^{\infty} B^{n-1-\nu} S(z_\nu). \tag{$\ddagger$}$$

The former is by induction and the first formula in ($\dagger$). The latter follows by letting m go to infinity in the formula

$$y_n = B^{-m} y_{n+m} - \sum_{\nu=n}^{n+m-1} B^{n-1-\nu} S(z_\nu)$$

which in turn follows by induction from the formula $y_n = B^{-1}y_{n+1} - B^{-1}S(z_n)$, a consequence of the second formula in ($\dagger$). The equations ($\ddagger$) may be written in the form

$$\Gamma(x, \gamma(x)) = 0, \tag{$*$}$$

where $\gamma(x)$ denotes the sequence $\{z_n\}$ and, for $x \in B_X$ and any sequence σ of elements of $B_{X\times Y}$ such that $\sigma_n \to 0$ as $n \to \infty$, $\Gamma(x, \sigma)$ is the sequence defined by

$$\Gamma(x, \sigma)_n = \sigma_n - \left(A^n x + \sum_{\nu=0}^{n-1} A^{n-1-\nu} R(\sigma_\nu), -\sum_{\nu=n}^{\infty} B^{n-1-\nu} S(\sigma_\nu) \right).$$

For a suitable Banach space S of sequences, $\Gamma\colon B_X \times S \to S$ is C^r and $\Gamma(0, 0) = 0$ and $D_2\Gamma(0, 0) = \text{identity}$. Thus by the implicit function theorem ($*$) can be solved and γ is C^r. It it easily shown that α is given by $\gamma(x)_0 = (x, \alpha(x))$, so that α is C^r by the smoothness of the evaluation map.

References

[1] R. Abraham & J. Robbin: *Transversal mappings and flows*, W. A, Benjamin, 1967.

[2] D. V. Anosov: *Geodesic flows on closed Riemann manifolds with negative curvature*, Proceedings of the Steklov Institute **90** (1967) AMS 1969.

[3] E. Coddington & N. Levinson, *Theory of ordinary differential equations*, McGraw-Hill, 1955.

[4] M. Hirsch, J. Palis, C. Pugh, M. Shub: Neighborhoods of Hyperbolic Sets, *Inv. Math.* **9**, 121-134 (1970)

[5] M. C. Irwin: On the stable manifold theorem. Bull. London Math. Soc. 2 1970 196-198.

[6] A. Kelley: The stable, center-stable, center, center-unstable manifolds, Appendix C in [1].

[7] O. E. Lanford, Lecture notes on dynamical systems, Part 2: Invariant manifolds, `http://www.math.ethz.ch/~lanford`

[8] J. Robbin, Stable Manifolds of Semi-Hyperbolic Fixed Points, Ill. J. of Math. (15), 595-609 (1970).

[9] S. J. van Strien, Center manifolds are not C^∞, Math. Z. **166** 143-145 (1979).