UC Berkeley

UC Berkeley Previously Published Works

Title

Stable manifolds and hyperbolic sets

Permalink

https://escholarship.org/uc/item/2rr3r1pz

Author

Hirsch, MW

Publication Date

1970

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

STABLE MANIFOLDS AND HYPERBOLIC SETS

MORRIS W. HIRSCH AND CHARLES C. PUGH

0. **Introduction.** Let U be an open set in a smooth manifold M and $f:U \to M$ a C^1 map. A fixed point x of f is *hyperbolic* if the derivative $T_xf:M_x \to M_x$ is an isomorphism and its spectrum is separated by the unit circle. If $T = T_xf$, this means that M_x has a unique splitting $E_1 \times E_2$ under T such that $T|E_1$ is expanding and $T|E_2$ is contracting. That is, for suitable equivalent norms on E_1 and E_2,

$$\max\{\|T^{-1}|E_1\|, \|T|E_2\|\} < 1.$$

The classical stable manifold theory says that this convenient behavior of T_xf is reflected in the behavior of f in a neighborhood V of x: there is a submanifold W^s of M tangent to E_2 at x such that

$$W^s \cap V = \{y \in V | \lim_{n\to\infty} (f|V)^n y = x\},$$

there is also a submanifold W^u tangent to E_1 such that

$$W^u \cap V = \{y \in V | \lim_{n\to\infty} (f|V)^{-n} y = x\}.$$

See for example Kelley [1, Appendix], [15] and [14], which contains further references.

We call W^s and W^u *local stable* and *unstable* manifolds of f at x, respectively. It turns out that they enjoy the same differentiability as f, and if f is C^k they depend continuously on f in the C^k topologies.

For technical reasons we allow M to be an infinite dimensional manifold modelled on a Banach space.

The notion of hyperbolic fixed point can be generalized to that of a *hyperbolic set* $\Lambda \subset U$. This means that $f(\Lambda) = \Lambda$, and $T_\Lambda M$ (the tangent bundle of M over Λ) has an invariant splitting $E_1 \oplus E_2$ such that $Tf|E_1$ is expanding and $Tf|E_2$ is contracting. (For this theory M is assumed finite dimensional and Λ compact, although generalizations are possible.) In Smale's theory of Ω-stability, and related topics [12], [13], "generalized stable manifold theorem" plays a key role: there is a neighborhood V of Λ, and submanifolds $W^s(x)$, $W^u(x)$ tangent to $E_2(x)$ and $E_1(x)$ respectively for each $x \in \Lambda$, such that

$$W^s(x) = \{y \in V | \lim_{n\to\infty} d((f|V)^n y, (f|V)^n x) = 0\},$$

$$W^u(x) = \{y \in V | \lim_{n\to\infty} d((f|V)^{-n} y, (f|V)^{-n} x) = 0\}.$$

If f is C^k, so are $W^s(x)$ and $W^u(x)$, and they depend continuously on f in the C^k topologies. Moreover $W^s(x)$ and $W^u(x)$ and their derivatives along $W^s(x)$ and $W^u(x)$ up to order k depend continuously on x.

The proof of the generalized stable manifold theorem proceeds in the following steps:

(A) Let $E = E_1 \times E_2$ be a Banach space; $T: E \to E$ a hyperbolic linear map expanding along E_1 and contracting along E_2; $E(r) \subset E$ the ball of radius r; and $f: E(r) \to E$ a Lipschitz perturbation of $T|E(r)$. The unstable manifold W for f will be the graph of a map $g: E_1(r) \to E_2(r)$ which satisfies $W = f(W) \cap E(r)$. We are led to consider the following transformation Γ_f in a suitable function space G of maps g:

$$\text{graph}\,\Gamma_f(g) = E(r) \cap f(\text{graph}\,g).$$

We call $\Gamma_f(g)$ the *graph transform* of g by f. The fixed point g_0 of Γ_f gives the unstable manifold of f; the existence of a fixed point is proved by the contracting map principle if f is sufficiently close to T pointwise, and the Lipschitz constant of $f - T$ is small enough.

(B) If f is C^k $(k \in Z_+)$ so is g_0. This is proved by induction on k. The successive approximations $\Gamma_f^n(g)$ converge C^k to g_0, but not, apparently, exponentially. The fibre contraction theorem (1.2) is needed to get this convergence.

(C) Let $\Lambda \subset U$ be a hyperbolic set. Let $\mathcal{M}$ be the Banach manifold of bounded maps $\Lambda \to M$, and $i \in \mathcal{M}$ the inclusion of Λ. Let $\mathcal{U} = \{h \in \mathcal{M} | h(\Lambda) \subset U\}$. Define $f_*: \mathcal{U} \to \mathcal{M}$ by

$$f_*(h) = f \circ h \circ f^{-1}.$$

Then f_* has a hyperbolic fixed point at i. By (A), f_* has a stable manifold $\mathcal{W}^s \subset \mathcal{M}$. For each $x \in \Lambda$, define $W^s(x) = ev_x(\mathcal{W}^s) = \{y \in M | y = \gamma(x) \text{ for some } \gamma \in \mathcal{W}^s\}$. It turns out that this definition gives a system of stable manifolds for f along Λ.

In §1 we collect various facts about maps and function spaces, including the Lipschitz inverse function theorem and the fiber contraction theorem. In §2 we carry out steps (A) and (B), and (C) is done in §3.

In §4 technical criteria for hyperbolicity are established. These are applied in §5 to prove that if $V \subset \Lambda$ is an invariant smooth submanifold then the non-wandering set of $f|V$ is a hyperbolic set for $f|V$.

The smoothness of the subbundles E^s and E^u, considered as fields of subspaces of $T_\Lambda M$, is studied in §6, which does not depend on the other sections. It is shown that E^s (also E^u) is always Hölder; if the stable manifolds have codimension 1, then E^s is Lipschitz, and is C^1 on every invariant C^2 submanifold of Λ if f is C^2. If the stable manifolds have codimension 1 and fill up an open set and f is C^2, then they provide a C^1 foliation of that set. Both the stable and the unstable foliations are C^1 if f is a C^2 volume preserving Anosov diffeomorphism of a compact 3-manifold.

§7 deals with hyperbolic sets of perturbations of f. There is a compact neighborhood V of Λ containing a unique maximal hyperbolic set Λ_0, and Λ_0 contains every invariant set in V. These maximal hyperbolic sets are structurally stable in a certain sense.

While preparing this paper we were fortunate to have had helpful discussions with participants in the 1968 Conference on Global Analysis in Berkeley. It is a pleasure to record our gratitude to them, especially to M. Shub, R. F. Williams and J. Moser. Special thanks are due to S. Smale for being the source of so many of the ideas developed in this paper.

1. **Lipschitz maps.** A map $f: X \to Y$ between metric spaces is *Lipschitz* if there exists a number k such that $d(fx, fy) \leq kd(x, y)$ for all $x, y \in X$. The smallest such k is the *Lipschitz constant* $L(f)$. If $X = Y$ and $L(f) < 1$ then f is a *contraction.* If f is not Lipschitz then $L(f) = \infty$.

1.1. (CONTRACTING MAP THEOREM). *Let X be a complete metric space and $f: X \to X$ a contraction. Then f has a unique fixed point x_0, equal to $\lim_{n\to\infty} f^n(x)$ for all $x \in X$. If $L(f) = k < 1$ and $g: X \to X$ has a fixed point y_0 such that $d(fy_0, gy_0) \leq \varepsilon$, then $d(x_0, y_0) \leq \varepsilon/(1-k)$.*

PROOF. The proof of the first assertion is quite standard and is therefore omitted. The last assertion is true because

$$\begin{aligned} d(x_0, y_0) &= d(fx_0, gy_0) \\ &\leq d(fx_0, fy_0) + d(fy_0, gy_0) \\ &\leq kd(x_0, y_0) + \varepsilon. \end{aligned}$$

If X is a set and Y a metric space, $\mathcal{M}(X, Y)$ denotes the space of all maps $X \to Y$ with the *uniform topology*, generated by the collection of sets

$$\{\mathcal{N}_\varepsilon(f) | f \in \mathcal{M}(X, Y), \varepsilon > 0\}$$

where

$$\mathcal{N}_\varepsilon(f) = \{g \in \mathcal{M}(X, Y) | d(fx, gx) < \varepsilon \text{ for all } x\}.$$

We write $d(f, g) = \sup_x\{d(fx, gx)\} \leq \infty$. If Y is complete and $f_0 \in \mathcal{M}(X, Y)$, the set of maps at a finite distance from f_0, namely

$$\{g \in \mathcal{M}(X, Y) | d(f_0, g) < \infty\} = \mathcal{M}(X, Y; f_0)$$

is a complete metric space with metric $d(f, g)$; it is open and closed in $\mathcal{M}(X, Y)$. A *bounded* subspace of $\mathcal{M}(X, Y)$ means a bounded subspace of some $\mathcal{M}(X, Y; f_0)$. If X is a space, the subset $C(X, Y) \subset \mathcal{M}(X, Y)$ of continuous maps is closed. If X is metric and $\lambda \geq 0$, then the subspace

$$\mathcal{H}_\lambda(X, Y) = \{f \in \mathcal{M}(X, Y) | L(f) \leq \lambda\}$$

is closed in $\mathcal{M}(X, Y)$.

The contracting map theorem has as a corollary: the function assigning to each contraction of a complete metric space its fixed point is continuous.

A fixed point x_0 of $f: X \to X$ is called *attractive* if $\lim_{n\to\infty} f^n(x) = x_0$ for all $x \in X$. The following extension of the contracting map principle is useful for proving maps to be C^1.

1.2. FIBER CONTRACTION THEOREM. *Let X be a space and $f: X \to X$ a map having an attractive fixed point $p \in X$. Let Y be a metric space and $\{g_x\}_{x \in X}$ a family of maps $g_x: Y \to Y$ such that the formula $F(x, y) = (fx, g_x y)$ defines a continuous map $F: X \times Y \to X \times Y$. Let $q \in Y$ be a fixed point for g_p. Then $(p, q) \in X \times Y$ is an attractive fixed point for F provided*

(a) $\limsup_{n \to \infty} L(g_{f^n x}) < 1$ *for each* $x \in X$.

PROOF. For each $(x, y) \in X \times Y$ we must prove $\lim_{n \to \infty} F^n(x, y) = (p, q)$. Therefore we could replace X by $\{p\} \bigcup \{f^n(x) | n \geq n_0\}$ for any n_0 and an arbitrary x. Hence we may assume instead, by (a), that

$$L(g_x) \leq \lambda < 1 \quad \text{for all } x \in X. \tag{1}$$

The theorem is proved if we show that $\pi_2 F^n(x, y) \to q$.

Call $d(g_{f^n x} q, q) = \delta_n$. Since $f^n(x) \to p$ as $n \to \infty$ and F is continuous at (p, q), $\delta_n \to 0$. By definition of F, $\pi_2 F^{n+1}(x, \cdot) = g_{f^n x} \circ \ldots \circ g_x$ and so

$$\begin{aligned} d(\pi_2 F^{n+1}(x, q), q) &\leq d(\pi_2 F^{n+1}(x, q), g_{f^n x}(q)) + d(g_{f^n x}(q), q) \\ &\leq \lambda d(\pi_2 F^n(x, q), q) + \delta_n \\ &\leq \lambda[\lambda d(\pi_2 F^{n-1}(x, q) + \delta_{n-1}] + \delta_n \\ &\leq \ldots \leq \lambda^n d(\pi_2 F(x, q), q) + \lambda^{n-1}\delta_1 + \ldots + \delta_0, \end{aligned}$$

which is $\sum_{j=0}^{n} \lambda^{n-j}\delta_j$. This tends to zero because, for any k, $0 < k < n$,

$$\begin{aligned} \sum_{j=0}^{n} \lambda^{n-j}\delta_j &= \sum_{j=0}^{k-1} \lambda^{n-j}\delta_j + \sum_{j=k}^{n} \lambda^{n-j}\delta_j \\ &\leq (\lambda^n + \ldots + \lambda^{n-k+1}) \max_j(\delta_j) + (1 + \ldots + \lambda^{n-k}) \max_{j \geq k}(\delta_j) \\ &\leq \lambda^{n-k} \max_j(\delta_j)/(1 - \lambda) + \max_{j \geq k}(\delta_j)/(1 - \lambda), \end{aligned}$$

which tends to zero as k and $n - k \to \infty$.

Hence

$$\begin{aligned} d(\pi_2 F^n(x, y), q) &\leq d(\pi_2 F^n(x, y), \pi_2 F^n(x, q)) \\ &+ d(\pi_2 F^n(x, q), q) \leq \lambda^n d(y, q) + \sum_{j=0}^{n-1} \lambda^{n-j}\delta_j \to 0 \quad \text{as } n \to \infty. \end{aligned}$$

REMARK 1. It is easy to see that if $A \subset X$ is such that $f(A) \subset A$, $f^n(x) \to p$ uniformly in A, and $L(g_x) \leq \kappa < 1$ for all $x \in A$, then the convergence $f^n(x, y)$ to (p, q) is uniform in $A \times Y$.

REMARK 2. Suppose f is a contraction of X and each g_x is a contraction of Y with $L(g_x)$ bounded away from 1. It is not clear whether there exists a metric on $X \times Y$ for which F is a contraction. The product metric will not suffice as the maps $f(x) = \frac{1}{2}x$, $g_x(y) = |x|^{1/2} + \frac{1}{2}(y - |x|^{1/2})$ show. This is the phenomenon of "shear" in the y-direction. It seems likely that no such metric exists in general.

Next we collect certain elementary relations between Lipschitz constants. In order to minimize notation, we assume the existence of all sums, compositions, inverses, etc. that are needed to make sense of the notation. Thus "$f + g$" implies that f and g are maps into a Banach space, and $f + g$ is the map $x \mapsto f(x) + g(x)$.

If X_1 and X_2 are metric spaces, the metric in $X_1 \times X_2$ is $d(x, y) = \sup\{d(x_1, y_1), d(x_2, y_2)\}$.

1.3. PROPOSITION.
(a) $L(f \circ g) \leq L(f)L(g)$.
(b) $L(f + g) \leq L(f) + L(g)$, *and* $L(f) - L(g) \leq L(f - g)$.
(c) $d(f_1 \circ g_1, f_2 \circ g_2) \leq d(f_1, f_2) + L(f_2)d(g_1, g_2)$.
(d) *Recall that* $\mathscr{H}_\lambda(Y, Z)$ *is the set of maps* $h: Y \to Z$ *with* $L(h) \leq \lambda$. *Composition defines a map* $\mathscr{H}_\lambda(Y, Z) \times C(X, Y) \to C(X, Z)$ *having Lipschitz constant* $\leq 1 + \lambda$.
(e) *For fixed* $h_0 \in \mathscr{H}_\lambda(Y, Z)$, *the map* $C(X, Y) \to C(X, Z)$ *given by* $g \mapsto h_0 \circ g$ *has Lipschitz constant* $\leq L(h_0)$.

PROOF. (a), (b) and (e) are obvious; (c) is proved by writing

$$d(f_1 \circ g_1, f_2 \circ g_2) \leq d(f_1 \circ g_1, f_2 \circ g_1) + d(f_2 \circ g_1, f_2 \circ g_2).$$

And (c) implies (d).

1.4. PROPOSITION. *Consider maps between topological subspaces of Banach spaces.*
(a) *If* f *is injective and* $L(f - g) < L(f^{-1})^{-1}$, *then* g *is injective, and* $L(g^{-1}) \leq [L(f^{-1})^{-1} - L(f - g)]^{-1}$.
(b) $|g^{-1} - h^{-1}| \leq L(g^{-1}) \cdot |h - g|$.
(c) *Let* $\mathscr{G}$ *be a space of invertible maps* g *such that* $L(g^{-1}) \leq \lambda$. *Then the Lipschitz constant of the map* $g \mapsto g^{-1}$ *is* $\leq \lambda$.

PROOF. (a): Follows from the two inequalities,

$$|gx - gy| \geq |fx - fy| - |(g - f)x - (g - f)y|$$

and

$$|fx - fy| \geq L(f^{-1})^{-1}|x - y|.$$

(b): Follows from

$$|g^{-1} - h^{-1}| = |g^{-1} \circ h \circ h^{-1} - g^{-1} \circ g \circ h^{-1}| \leq L(g^{-1}) \cdot |h - g|.$$

This proves (c).

The standard modern proof of the Inverse Function Theorem (see [9]) deals with a C^1 small perturbation of an invertible linear map; a C^1 inverse is produced. Abstracting this idea leads to the following result.

1.5. LIPSCHITZ INVERSE FUNCTION THEOREM. *Let* E, F *be Banach spaces,* $U \subset E$ *and* $V \subset F$ *open sets and* $f: U \to V$ *a homeomorphism such that* f^{-1} *is Lipschitz. Let* $h: U \to F$ *be such that* $L(h)L(f^{-1}) < 1$ *and put* $g = f + h: U \to F$. *Then* g *is a homeomorphism onto an open set, and*

$$L(g^{-1}) \leq L(f^{-1})/(1 - L(g - f)L(f^{-1})) = [L(f^{-1})^{-1} - L(g - f)]^{-1}.$$

Note that no assumption on the existence of a linear isomorphism $E \to F$ (to which f might be nearly tangent) is made. In fact it is not clear that a lipeomorphism between open subsets of E and F implies the existence of a linear isomorphism between E and F.

PROOF. The injectivity of g and the estimate on $L(g^{-1})$ follow from 1.4a. Since g^{-1} is Lipschitz, it is continuous. It remains to prove that if $x_0 \in U$, then g maps some neighborhood of x_0 onto a neighborhood of $g(x_0)$. We may assume that $h(x_0) = 0$ and $g(x_0) = f(x_0)$; if not, replace h by the map $x \mapsto h(x) - h(x_0)$.

Let juxtaposition denote composition, and put $hf^{-1} = v: V \to F$. We shall prove that $I + v: V \to F$ sends a neighborhood N of $f(x_0)$ onto an open set N'; then $f + h = (f + h)f^{-1}f = (I + v)f$ maps $f^{-1}N$ onto $(I + v)N = N'$.

We may assume $f(x_0) = 0 \in F$, and hence $v(0) = 0$. Put $L(v) = \lambda < 1$. Let V contain $B_r = B_r(0)$, the ball in F of radius r and center 0. Put $s = r(1 - \lambda)$ and let $\mathscr{X}$ be the complete metric space of maps $w: B_s \to F$ such that $w(0) = 0$ and $L(w) \leq \lambda/(1 - \lambda)$. We seek a map $w_0 \in \mathscr{X}$ such that $I + w_0$ is a right inverse for $I + v$; this is equivalent to $w_0 = -v(I + w_0)$.

If $w \in \mathscr{X}$ then $(I + w)B_s \subset B_r$, since if $|x| \leq s$, we have

$$|(I + w)x| \leq |x| + L(w)|x| \leq s(1 + \lambda/(1 - \lambda)) = r.$$

Therefore the composition $-v(I + w) = \Phi(w)$ is defined; it is easy to compute that $\Phi(w) \in \mathscr{X}$. The map $\Phi: \mathscr{X} \to \mathscr{X}$ has Lipschitz constant $\leq L(v) < 1$, so that Φ has a unique fixed point $w_0 \in \mathscr{X}$. Since $(I + v)(I + w)x = x$ for all $x \in B_s$, it follows that $I + v$ maps the open set $(I + v)^{-1}$ (int B_s) onto int B_s. This completes the proof of 1.5.

1.6. SIZE ESTIMATE. *Let X, Y be metric spaces and $f: X \to Y$ a bijective map such that $L(f^{-1})^{-1} \geq \lambda$. Then $fB_r(x) \supset B_{\lambda r}(fx)$ for all $r > 0$ and $x \in X$.*

PROOF. If $d(y, x) > r$, then $r < d(y, x) \leq \lambda^{-1}d(fy, fx)$, so that $d(fy, fx) > dr$. This means $f(X - B_r(x)) \subset Y - B_{\lambda r}(fx)$. Since f is bijective the result follows.

REMARK 1. The Lipschitz I.F.T. (1.5) asserts that $|L(g^{-1}) - L(f^{-1})| \to 0$ with $L(g - f)$, but says nothing about $L(g^{-1} - f^{-1})$. If f is a diffeomorphism then $L(g^{-1} - f^{-1})$ does approach 0 with $L(g - f)$. Consider however the nondifferentiable homeomorphism $f: R \to R$ given by

$$f(x) = x \quad \text{if } x \leq 0,$$
$$\quad = 2x \quad \text{if } x \geq 0.$$

Let $g(x) = f(x) + \varepsilon$. Then $L(f - g) = 0$, but if $\varepsilon \neq 0$ then $L(f^{-1} - g^{-1}) \geq \frac{1}{2}$.

REMARK 2. The Size Estimate can be extended to local homeomorphisms in Banach spaces E, F as follows.

1.7. PROPOSITION. *Let $U \subset E$ be open and $f: U \to F$ a local homeomorphism whose local inverses all have Lipschitz constants $\leq \lambda^{-1}$. If $B_r(x) \subset U$, then $B_{\lambda r}(fx) \subset fB_r(x)$ and there is a unique continuous map $g: B_{\lambda r}(fx) \to B_r(x)$ such that $gf(x) = x$ and $f \circ g = I$. Moreover $L(g) \leq \lambda^{-1}$.*

PROOF. Left to the reader.

2. The invariant manifolds of a hyperbolic fixed point. Let E be a Banach space and T an isomorphism of E; that is, a linear homeomorphism of E onto E. We call

T hyperbolic if its spectrum lies off the unit circle. If T is hyperbolic so is T^{-1} because the spectrum of T^{-1} is the set of reciprocals of elements of the spectrum of T.

If T is hyperbolic there is a unique splitting $E_1 \times E_2 = E$ invariant under T such that the spectrum of $T_1 = T|E_1$ lies outside the unit circle, while that of E_2 is inside; see [11]. Moreover E_1 and E_2 can be renormed so that $\|T_2\| < 1$ and $\|T_1^{-1}\| < 1$; see [8], and also Theorem 3.1 below. We shall always assume that E_1 and E_2 have such norms, and that on E the norm is $|(x_1, x_2)| = \max\{|x_1|, |x_2|\}$, for $x_1 \in E_1$ and $x_2 \in E_2$.

The quantity $\tau = \max(\|T_2\|, \|T_1^{-1}\|) < 1$ is quite useful. We shall call it the *skewness* of T.

It is convenient to put $m(T_1) = \|T_1^{-1}\|^{-1}$. Then $|T_1 x| \geq m(T_1)|x|$.

The iterative behavior of a hyperbolic T is described by the following result.

2.0. PROPOSITION. *Let $T: E \to E$ be a hyperbolic linear map and let $E = E_1 \times E_2$ canonically. The points of E_1, E_2 are characterized respectively by $|T^{-n}x| \to 0$, $|T^n x| \to 0$, or equivalently by these quantities being bounded, as $n \to \infty$. Furthermore, if V is a bounded neighborhood of 0 in E_1 such that $TV \supset V$ then $(\bigcup_{n\geq 0} T^{-n}W) \cup E_2$ is a neighborhood of E_2 for any uniform neighborhood W of $TV - V$ in E.*

PROOF. A uniform neighborhood of a set A in a metric space X is any set containing $N_\varepsilon(A) = \{x \in X : d(x, a) \leq \varepsilon \text{ for some } a \in A\}$ for some $\varepsilon > 0$.

The first statement is trivial. The second statement is proved as follows.

Since V is bounded there is, for each $x_1 \in E_1$, $x_1 \neq 0$, a largest value $n = n(x_1)$ such that $T^n x_1 \in V$. Thus, $T^{n+1} x_1 \in TV - V$. Clearly, $n \leq \log(\text{radius } V/|x_1|)/\log(m(T_1))$ since $m(T_1)^n |x_1|$ must be no greater than radius V.

Hence, for any $(x_1, x_2) \in E$ with $x_1 \neq 0$, $T^n(x_1, x_2) \in N_\varepsilon(TV - V)$ for $n = n(x_1)$ and $\varepsilon = \|T_2\|^n |x_2|$. This proves 2.0. In fact this proves $(\bigcup_{n\geq 0} T^{-n}W) \cup E_2$ to be a neighborhood of E_2 which is nonuniform only at infinity. Q.E.D.

The object of stable manifold theory is to demonstrate similar behavior for suitable perturbations of T, and for the more general situation where the fixed point is replaced by a hyperbolic set.

Throughout the rest of §2, T will be a hyperbolic isomorphism $E \to E$ of skewness τ, $0 < \tau < 1$.

Let $E(r) \subset E$ be the closed ball of radius r about 0.

We shall be interested in proving that for $f: E(r) \to E$, close enough to T,

$$W_1 = \bigcap_{n\geq 0} f^n(E(r)), \qquad W_2 = \bigcap_{n\geq 0} f^{-n}(E(r)),$$

are submanifolds close to $E_1(r)$, $E_2(r)$. We shall assume $r < \infty$.

2.1. LEMMA. *Let $f: E(r) \to E$ satisfy $L(f - T) < \varepsilon < (1 - \tau)/(1 + \tau)$. If $x, y \in E(r)$ and $|x_1 - y_1| \geq |x_2 - y_2|$ then*

$$\begin{aligned} |f_1(x) - f_1(y)| &\geq (\tau^{-1} - \varepsilon)|x_1 - y_1| \\ &\geq (\tau + \varepsilon)|x_1 - y_1| \\ &\geq |f_2(x) - f_2(y)|. \end{aligned}$$

PROOF.
$$\begin{aligned}|f_1(x)-f_1(y)| &= |T_1(x-y)+(f_1-T_1)(x)-(f_1-T_1)(y)|\\ &\geq |T_1(x_1-y_1)|-\varepsilon|x-y|\\ &\geq \tau^{-1}|x_1-y_1|-\varepsilon|x-y|\\ &= (\tau^{-1}-\varepsilon)|x_1-y_1|,\end{aligned}$$

since $|x-y| = \max(|x_1-y_1|, |x_2-y_2|) = |x_1-y_1|$. Similarly, $|f_2(x)-f_2(y)| \leq \tau|x_2-y_2| + \varepsilon|x-y| \leq (\tau+\varepsilon)|x_1-y_1|$. Since $\varepsilon < (1-\tau)/(1+\tau)$ we have $\varepsilon < 1-\tau$ and $\varepsilon < \tau^{-1}-1$ which shows that $\tau+\varepsilon < 1 < \tau^{-1}-\varepsilon$, completing the proof. Q.E.D.

2.2. PROPOSITION. *If $f: E(r) \to E$ and $L(f-T) < \varepsilon < (1-\tau)/(1+\tau)$, then $W_1 = \bigcap_{n\geq 0} f^n(E(r))$ is the graph of a function $U_1 \to E_2(r)$ while $W_2 = \bigcap_{n\geq 0} f^{-n}(E(r))$ is the graph of a function $U_2 \to E_1(r)$, where U_1, U_2 are subsets of $E_1(r)$, $E_2(r)$.*

REMARKS. This permits U_1 or U_2 to be empty. The notation $f^{-n}(E(r))$ means $\{x \in E(r): f(x), \ldots, f^n(x) \text{ are defined and in } E(r)\}$. One should regard 2.2 as a uniqueness theorem.

PROOF. We deal first with W_2. Let $x, y \in W_2$ with $x_2 = y_2$. We must show $x_1 = y_1$. By assumption $f^n(x), f^n(y) \in E(r)$ for all $n \geq 0$. By 2.1,

$$\begin{aligned}2r \geq |(f^n)_1 x - (f^n)_1 y| &= |f_1(f^{n-1}x) - f_1(f^{n-1}y)|\\ &\geq (\tau^{-1}-\varepsilon)|(f^{n-1})_1 x - (f^{n-1})_1 y| \geq \cdots \geq (\tau^{-1}-\varepsilon)^n|x_1-y_1|.\end{aligned}$$

As $n \to \infty$, $(\tau^{-1}-\varepsilon)^n \to \infty$ and so $x_1 = y_1$. Hence W_2 is a graph as claimed.

Now suppose $x, y \in W_1$ and $x_1 = y_1$. We must show $x_2 = y_2$. For every $n \geq 0$ there exist $x', y' \in E(r)$ such that $f^n(x') = x, f^n(y') = y$. Suppose, for some $0 \leq j < n$,

$$|(f^j x')_1 - (f^j y')_1| \geq |(f^j x')_2 - (f^j y')_2|.$$

Then by repeated application of 2.1,

$$|(f^n x')_1 - (f^n y')_1| \geq |(f^n x')_2 - (f^n y')_2|,$$

which is the same as $0 \geq |x_2 - y_2|$, so that $x_2 = y_2$.

On the other hand if

$$|(f^j x')_1 - (f^j y')_1| \leq |(f^j x')_2 - (f^j y')_2|$$

for all $0 \leq j \leq n$ then

$$\begin{aligned}|x_2-y_2| &= |(f^n x')_2 - (f^n y')_2|\\ &= |f_2(f^{n-1}x') - f_2(f^{n-1}y')|\\ &\leq \tau|(f^{n-1}x')_2 - (f^{n-1}y')_2| + \varepsilon|f^{n-1}(x') - f^{n-1}(y')|\\ &= (\tau+\varepsilon)|(f^{n-1}x')_2 - (f^{n-1}y')_2|\\ &\leq \cdots \leq (\tau+\varepsilon)^n|x'_2 - y'_2|\\ &\leq 2r(\tau+\varepsilon)^n.\end{aligned}$$

Since $|x_2 - y_2|$ is independent of n and $(\tau+\varepsilon)^n \to 0$, we have $x_2 = y_2$. This shows that W_1 is a graph as claimed, proving 2.2.

Alternately, we could have proved the second part of 2.2 by looking at W_2 for

f^{-1}. Unfortunately, f^{-1} is not defined on $E(r)$, so it was easier to proceed directly. We shall come again to this problem in 2.4.

DEFINITION. A *stable manifold* W^s for $f: E(r) \to E$ is the set of $x \in E(r)$, such that $|f^n x|$ is defined and stays bounded as $n \to \infty$. Clearly $fW^s \subset W^s$. An *unstable manifold* W^u for $f: E(r) \to E$ is the set of $x \in E(r)$ such that $|f^{-n}x|$ is defined and stays bounded as $n \to \infty$. Clearly $f^{-1}W^u \subset W^u$.

For suitable f, it turns out that $W^u = W_1$, $W^s = W_2$, and W^s is contracted toward the point $W^u \cap W^s$ by f while W^u is contracted toward the point $W^u \cap W^s$ by f^{-1}.

Before stating the basic theorem for unstable manifolds, we define the *graph transform.* If $E = E_1 \times E_2$, $r > 0$, and $f: E(r) \to E$, then we write $\Gamma_f(g) = h$ provided $h, g: E_1(r) \to E_2(r)$ and

$$f(\text{graph}(g)) \cap E(r) = \text{graph}(h).$$

Γ_f is called the graph transform for f. If $T = T_1 \times T_2$ is hyperbolic respecting $E_1 \times E_2$ then for any $g: E_1(r) \to E_2(r)$, $\Gamma_T(g)$ is defined and equals $T_2 \circ g \circ T_1^{-1}|E_1(r)$. Similarly, if $\Gamma_f(g)$ exists then $\Gamma_f(g) \circ f_1 \circ (1, g) = f_2 \circ (1, g)$ where 1 is the identity map of $E_1(r)$. If $f_1 \circ (1, g)$ is invertible this means that

$$\Gamma_f(g) = f_2 \circ (1, g) \circ [f_1 \circ (1, g)]^{-1}|E_1(r).$$

DEFINITION. $C^k(X, Y)$ is the set of functions $f: X \to Y$ of class C^k (first k derivatives exist at all points of X and are continuous) with bounded k norm

$$|f|_k = \sup_{x \in X} \max(|f(x)|, \|(Df)_x\|, \dots, \|(D^k f)_x\|).$$

DEFINITION. For $r > 0$ fixed, we let $\mathcal{N}_\varepsilon^k(T) = \{f \in C^k(E(r), E): L(f - T) < \varepsilon$ and $|f(0)| < \varepsilon\}$. This could be called a "Lipschitz neighborhood of $T|E(r)$ in $C^k(E(r), E)$."

The k-norm makes $C^k(E(r), E)$ complete and defines the so-called *uniform C^k topology.*

2.3. UNSTABLE MANIFOLD THEOREM FOR A POINT. *Given $0 < \tau < 1$ and $r > 0$ there exist $\varepsilon > 0$, independent of r, and $0 < \delta < \varepsilon$ with the following properties. If $T = T_1 \times T_2$ is a hyperbolic linear operator on $E_1 \times E_2$ of skewness τ and $f: E(r) \to E$ is a Lipschitz map satisfying $L(f - T) < \varepsilon$, $|f(0)| \le \delta$, then there is a unique map $g_f: E_1(r) \to E_2(r)$ whose graph is $W_1 = \bigcap_{n \ge 0} f^n(E(r))$. Moreover $L(g_f) < 1$ and g_f is of class C^k iff f is. The assignment $f \mapsto g_f$ is continuous as a map $\mathcal{N}_\delta^k(T) \to C^k(E_1(r), E_2(r))$. The map $(f|W_1)^{-1}: W_1 \to W_1$ is a contraction of W_1 into its interior (that is, into $\{x \in W_1: |x_1| \le s\}$ for some $s < r$).*

REMARK 1. By 2.2 uniqueness of g_f is assured; it is only a matter of producing a function $E_1(r) \to E_2(r)$ whose graph is contained in its own f-image.

REMARK 2. Choice of ε and δ are restricted no more than $\varepsilon < (1 - \tau)/(1 + \tau)$, $\delta < \varepsilon^2 r\tau$.

PROOF OF THEOREM 2.3. We shall first prove 2.3 under the assumption $f(0) = 0$. Afterwards the general case is easily handled.

Choose any

$$0 < \varepsilon < (1 - \tau)/(1 + \tau) = (\tau^{-1} - 1)/(\tau^{-1} + 1).$$

We claim this ε does the job; three forms of the inequality are used:

$$\varepsilon < 1 - \tau, \varepsilon < \tau^{-1}, \text{and } (\tau + \varepsilon)/(1 - \tau\varepsilon) < 1.$$

Let $\mathscr{G} = \{g \in \mathscr{M}(E_1(r), E_2(r)) \colon g(0) = 0 \text{ and } L(g) \leq 1\}$. We show that $\Gamma_f : \mathscr{G} \to \mathscr{G}$ is a well-defined contraction. Clearly $\mathscr{G}$ is complete. Put

$$\psi_f(g) = f_1 \circ (1, g), \qquad \phi_f(g) = f_2 \circ (1, g).$$

Thus $\psi_f(g)$ is a map $E_1(r) \to E_1$ and $\phi_f(g)$ is a map $E_1(r) \to E_2$. Restricting to $E_1(r)$, we have $L(\psi_f(g) - T_1) \leq L(f - T) < \varepsilon$. Hence, by the Lipschitz inverse function theorem (1.5), $\psi_f(g)$ is a lipeomorphism, and

$$L([\psi_f(g)]^{-1}) \leq [L(T_1^{-1})^{-1} - L(\psi_f g - T_1)]^{-1} \leq [\tau^{-1} - \varepsilon]^{-1}$$

and so, since $(\psi_f g)(0) = 0$,

$$(\psi_f g)(E_1(r)) \supset E_1(r(\tau^{-1} - \varepsilon)) \supset E_1(r).$$

Hence $[\psi_f g]^{-1} | E_1(r)$ is a well-defined map into $E_1(r)$ with Lipschitz constant $\leq [\tau^{-1} - \varepsilon]^{-1} < 1$. Since $L(\phi_f g) \leq (\tau + \varepsilon) < 1$ and $(\phi_f g)(0) = 0$, it follows that $\Gamma_f(g) = \phi_f g \circ [\psi_f g]^{-1} | E_1(r)$ is a well defined map $\mathscr{G} \to \mathscr{G}$.

To see that Γ_f is a contraction, take $g_1, g_2 \in \mathscr{G}$ and estimate

$$\begin{aligned}
|\Gamma_f(g_1) - \Gamma_f(g_2)| &\leq |(\phi_f g_1) \circ (\psi_f g_1)^{-1} - (\phi_f g_1) \circ (\psi_f g_2)^{-1}| \\
&\quad + |(\phi_f g_1) \circ (\psi_f g_2)^{-1} - (\phi_f g_2) \circ (\psi_f g_2)^{-1}| \\
&\leq L(\phi_f g_1) |(\psi_f g_1)^{-1} - (\psi_f g_2)^{-1}| + |\phi_f g_1 - \phi_f g_2| \\
\text{(by 1.4(b))} \quad &\leq (\tau + \varepsilon) L((\psi_f g_1)^{-1}) \cdot |\psi_f g_1 - \psi_f g_2| \\
&\quad + (\tau + \varepsilon)|g_1 - g_2| \\
&\leq (\tau + \varepsilon)(\tau^{-1} - \varepsilon)^{-1} \varepsilon |g_1 - g_2| + (\tau + \varepsilon)|g_1 - g_2| \\
&= ((\tau + \varepsilon)/(\tau^{-1} - \varepsilon))(\varepsilon + (\tau^{-1} - \varepsilon))|g_1 - g_2| \\
&= ((\tau + \varepsilon)/(1 - \tau\varepsilon))|g_1 - g_2|.
\end{aligned}$$

Hence Γ_f has a unique fixed point $g_f \in \mathscr{G}$.

As remarked before, 2.2 now implies that $W_1 = \bigcap_{n \geq 0} f^n(E(r)) = \text{graph } g_f$.

Consider the following commutative diagram

$$\begin{array}{ccc}
W_1 & \xrightarrow{\quad f \quad} & f(W_1) \\
\pi_1 \downarrow & & \downarrow \pi_1 \\
E_1(r) & \xrightarrow[f_1 \circ (1, g_f)]{} & E_1
\end{array}$$

where $\pi_1(x_1, x_2) = x_1$. Since $L(g_f) \leq 1$, we have $|x - y| = |x_1 - y_1|$ for $x = (x_1, x_2)$, $y = (y_1, y_2) \in W_1$. This shows that π_1 is an isometry from W_1 onto $E_1(r)$. Since

$f_1 \circ (1, g) = \psi_f(g_f)^{-1}$ is a contraction, the same is true for $f|W_1$. Therefore $f|W_1$ has the unique fixed point 0. Any fixed point of f in $E(r)$ must belong to the unstable manifold W_1, and so 0 is the only fixed point of f.

Next we investigate the differentiability of g_f. Assume $f \in \mathcal{N}^1_\varepsilon(T)$, with ε as above, and $f(0) = 0$. Define

$$\Delta f:(E \times E)(r) \to E \times E,$$
$$(x, y) \mapsto (fx, Df_x y);$$

similarly for $g \in \mathcal{G}^1 = \mathcal{G} \cap C^1(E_1(r), E_2(r))$,

$$\Delta g: E_1(r) \times E_1(r) \to E_2(r) \times E_2(r),$$
$$(x_1, y_1) \mapsto (gx_1, (Dg)_{x_1} y_1).$$

Even though Δf is not Lipschitz close to $T \times T$, we still claim the graph transform for Δf is a well defined fiber contraction of $\mathcal{G} \times \mathcal{H}$, where

$$\mathcal{H} = \{h \in C^0(E_1(r), L(E_1, E_2)| \; \|h(x_1)\| \le 1 \text{ for all } x_1 \in E_1\}$$
$$= \text{unit ball in the Banach space of continuous bounded maps } E_1(r) \to L(E_1, E_2).$$

Indeed, for $(g, h) \in \mathcal{G} \times \mathcal{H}$ we put

$$\Gamma_{\Delta f}(g, h) = (\Gamma_f g(x_1), (Df_2)_\xi \circ (1, h(\xi_1)) \circ [(Df_1)_\xi \circ (1, h(\xi_1))]^{-1})$$

where $\xi_1 = (\psi_f g)^{-1}(x_1)$, $\xi_2 = g(\xi_1)$, and $\xi = (\xi_1, \xi_2)$. One can see directly that $\Gamma_{\Delta f}$ is well defined and contracts fibers; but it is instructive to be more formal.

Composition on the left by T defines a hyperbolic map of skewness τ, namely $T_\#:\mathcal{E}(1) \to \mathcal{E}$ for $\mathcal{E} = L(E, E)$; the canonical invariant splitting $\mathcal{E} = \mathcal{E}_1 \times \mathcal{E}_2$ is obtained by setting $\mathcal{E}_i = L(E, E_i)$. (Recall that $\mathcal{E}(1)$ is the unit ball in $\mathcal{E}$.)

If $\|T' - T\| \le \varepsilon$ then composition on the left by T' defines $T'_\#:\mathcal{E}(1) \to \mathcal{E}$ and $L(T'_\# - T_\#) = L(T' - T) \le \varepsilon$. Hence $\Gamma_{T'_\#}$, the graph transform for $T'_\#$, is a contraction of

$$\mathcal{G}\mathcal{E} = \{H \in C^0(\mathcal{E}_1(1), \mathcal{E}_2(1)) | H(0) = 0 \text{ and } L(H) \le 1\},$$

well defined by the formula

$$\Gamma_{T'_\#}(H)(S_1) = T'_2 \circ (1, H(S_1)) \circ [T'_1 \circ (1, H(S_1))]^{-1}.$$

For any $h \in \mathcal{H}$ and $\xi_1 \in E_1(r)$ we apply the functor $(\cdot)_\#$ to the linear map $h(\xi_1): E_1 \to E_2$ to get $h(\xi_1)_\# \in \mathcal{G}\mathcal{E}$. The definition of $\Gamma_{\Delta f}$ can be written

$$\Gamma_{\Delta f}(g, h)(x_1) = ((\Gamma_f g)(x_1), \Gamma_{(Df_\xi)_\#}(h(\xi_1)_\#)).$$

It is now clear that $\Gamma_{\Delta f}:\mathcal{G} \times \mathcal{H} \to \mathcal{G} \times \mathcal{H}$ contracts both base and fibers uniformly. Therefore, by the fiber contraction theorem (1.2) there is a unique attractive fixed point (g_f, h_f) of $\Gamma_{\Delta f}$.

Observe that $\Delta(\Gamma_f g) = \Gamma_{\Delta f}(\Delta g)$; this is a consequence of the naturality of Δ, which is just the tangent functor. Take, then, $g_0 \equiv 0$ and consider the convergence $(\Gamma_{\Delta f})^n(\Delta g_0) \to (g_f, h_f)$ as $n \to \infty$. But $(\Gamma_{\Delta f})^n(\Delta g_0) = \Delta(\Gamma^n_f f_0)$, so that $\Delta(\Gamma^n_f g_0)$ converges to (g_f, h_f); thus $D(g_f)$ exists and equals h_f.

Now suppose f is of class C^k, $k \geq 2$, $|f|_k \leq M$, and 2.3 holds for $k - 1 \geq 1$. Choose ε', $\varepsilon < \varepsilon' < (1 - \tau)/(1 + \tau)$, and choose r', $0 < r' \leq r$, so that $Mr' < \varepsilon' - \varepsilon$, $r' \leq 1$. Then let $f' = f|E(r')$. We claim that $\Delta f' \in \mathcal{N}_{\varepsilon'}(T \times T)$. We have $L(\Delta f' - T \times T) < \varepsilon'$ since

$$\begin{aligned}
&|(\Delta f' - T \times T)(x, y) - (\Delta f' - T \times T)(x', y')| \\
&\leq \max(|(f' - T)(x) - (f' - T)(x')|, |((Df')_x - T)y - ((Df')_{x'} - T)y'|) \\
&\leq \max(\varepsilon|x - x'|, |[(Df')_x - (Df')_{x'}]y| + |[(Df')_{x'} - T](y - y')|) \\
&\leq \max(\varepsilon|x - x'|, M|x - x'|r' + \varepsilon|y - y'|) \\
&< \varepsilon' \max(|x - x'|, |y - y'|),
\end{aligned}$$

for $|x - x'| \geq |y - y'|$ implies

$$Mr'|x - x'| + \varepsilon|y - y'| \leq (Mr' + \varepsilon)|x - x'| < \varepsilon'|x - x'|,$$

while $|y - y'| \geq |x - x'|$ implies

$$Mr'|x - x'| + \varepsilon|y - y'| < (Mr' + \varepsilon)|y - y'|.$$

Hence $\Delta f' \in \mathcal{N}_{\varepsilon'}(T \times T)$ so by the induction hypotheses there is a unique solution $g_{\Delta f'}$ of $\Gamma_{\Delta f'}(g_{\Delta f'}) = g_{\Delta f'}$ and $g_{\Delta f'}$ is of class C^{k-1}. But clearly, $g_{\Delta f}|E_1(r') \times E_1(r')$ solves this equation too. Hence g_f is of class C^k on $E_1(r')$.

We know that $(f|W)^{-1}: W_1 \to W_1$ is a contraction with $L((f|W_1)^{-1}) \leq (\tau^{-1} - \varepsilon)^{-1}$. So choosing $N > \log(r/r')/\log(\tau^{-1} - \varepsilon)$ we have

$$\pi_1(f|W)^{-N}(E_1(r)) \subset E(r \cdot (\tau^{-1} - \varepsilon)^{-N}) \subset E(r').$$

Since $\Gamma_f(g_f) = g_f$ we have

$$(*) \qquad g_f = (f^N)_2 \circ (1, g_{f'}) \circ [f^N \circ (1, g_{f'})]^{-1}|E_1(r).$$

Hence g_f is of class C^k on all $E_1(r)$. This formula (*) holds for any $f \in \mathcal{N}_\varepsilon(T)$ with the same N. Hence as $\bar{f} \to f$ in $\mathcal{N}^k_\varepsilon(T)$, it is clear that $\bar{f}' \to f'$, C^k, which implies $\Delta \bar{f}' \to \Delta f'$, C^{k-1}, which implies by induction (since $\Delta f' \in \mathcal{N}_\varepsilon(T \times T)$) $g_{\Delta \bar{f}'} \to g_{\Delta f'}$, C^{k-1}, which implies $g_{\bar{f}'} \to g_{f'}$, C^k. Hence, using the formula (*) for $g_{\bar{f}}$, we have $g_{\bar{f}} \to g_f$, C^k, as $\bar{f} \to f$, C^k.

This completes the proof of 2.3 in case $f(0) = 0$.

The case $f(0) \neq 0$ is obtained formally; the idea is to translate the origin of E to the fixed point of f, which we must prove exists. We may assume 2.3 true for all $0 < \tau < 1$ and $r > 0$ provided the map involved takes 0 to 0 and lies in $\mathcal{N}_\varepsilon(T)$, $\varepsilon < (1 - \tau)/(1 + \tau)$.

Choose $0 < \delta < \min(\varepsilon, \varepsilon^2 r\tau)$. For $|f(0)| < \delta$ and $L(f - T) < \varepsilon$, f has a unique fixed point $p_f = (p_1, p_2)$ in $E(r)$. To see this observe that the map

$$\bar{f}: (x_1, x_2) \mapsto (T_1^{-1}(x_1 - T_1 x_1 - f_1(x_1, x_2)), f_2(x_1, x_2))$$

is a contraction of $E(r)$ and has the same fixed points as f. By the contracting map theorem (1.1) p_f depends continuously on $f \in \mathcal{N}_\delta(T)$, and $|p_f| \leq \delta(1 - \tau - \varepsilon)^{-1}$.

We remark that the inequality $\delta < \varepsilon^2 r\tau$ implies

$$\delta < \frac{(\tau^{-1} - \varepsilon - 1)(1 - \tau - \varepsilon)r}{1 + \tau^{-1} - \varepsilon},$$

for

$$\varepsilon^2 r\tau < \varepsilon^2 r\tau(1 + \tau) < \left(\frac{1 - \tau}{1 + \tau}\right)^2 r\tau(1 + \tau)$$

$$= \frac{(1 - \tau)(1 - \tau)r}{\tau^{-1} + 1} = \frac{(\tau^{-1} - 1)(\tau - \tau^2)r}{1 + \tau^{-1}}$$

$$= \frac{\left(\tau^{-1} - \dfrac{1 - \tau}{1 + \tau} - 1\right)\left(1 - \tau - \dfrac{1 - \tau}{1 + \tau}\right)r}{1 + \tau^{-1}}$$

$$< \frac{(\tau^{-1} - \varepsilon - 1)(1 - \tau - \varepsilon)r}{1 + \tau^{-1} - \varepsilon}.$$

Put $\hat{r} = r - \delta(1 - \tau - \varepsilon)^{-1} \geq r - |p_f|$. Define

$$\hat{f}: E(\hat{r}) \to E, \qquad \hat{f}(x) = f(x + p_f) - p_f.$$

Observe that $\hat{f}$ is well defined, $L(\hat{f} - T) = L(f - T) < \varepsilon$, and $\hat{f}(0) = 0$. Hence there exists a unique $\hat{g} = g_{\hat{f}}: E_1(\hat{r}) \to E_2(\hat{r})$ such that $W_1(\hat{f}) = \bigcap_{n \geq 0} \hat{f}^n E(\hat{r}) = \text{graph}(g_{\hat{f}})$. Define $g = g_f: E_1(r) \to E_2(r)$,

$$g(x_1) = p_2 + \phi_{\hat{f}}(g_{\hat{f}}) \circ [\psi_{\hat{f}}(g_{\hat{f}})]^{-1}(x_1 - p_1).$$

To see that g is well defined, observe that $\psi_{\hat{f}}(\hat{g})E_1(\hat{r}) \supset E_1((\tau^{-1} - \varepsilon)\hat{r})$, since $L(\psi_{\hat{f}}(\hat{g})^{-1}) \leq (\tau^{-1} - \varepsilon)^{-1}$, as was shown earlier. Moreover the estimate for δ and the definition of $\hat{r}$ imply $(\tau^{-1} - \varepsilon)\hat{r} \geq r + \delta(1 - \tau - \varepsilon)^{-1}$. This implies that $[\psi_{\hat{f}}(\hat{g})]^{-1}$ is defined on $x_1 - p_1$ if $|x_1| \leq r$, and also that $g(E_1(r)) \subset E_2(r)$.

Clearly $L(g) = L(\phi_{\hat{f}}(\hat{g}) \circ [\psi_{\hat{f}}(\hat{g})]^{-1}) = L(\hat{g}) \leq 1$; and $g(p_1) = p_2$.

We claim $f(\text{graph}\, g) \cap E(r) = \text{graph}(g)$. Let $x_1 \in E_1(r)$ and assume $f_1(x_1, g(x_1)) \in E_1(r)$. Then $|x_1 - p_1| < \hat{r}$, for by the Lipschitz I.F.T. (1.5),

$$|f_1(x_1, g(x_1)) - p_1| \geq (\tau^{-1} - \varepsilon)|x_1 - p_1|,$$

which implies

$$|x_1 - p_1| \leq \frac{r + |p_1|}{\tau^{-1} - \varepsilon} \leq r + \frac{\delta(1 - \tau - \varepsilon)^{-1}}{\tau^{-1} - \varepsilon} \leq \hat{r},$$

by our choices of δ and $\hat{r}$. Therefore $x_1 - p_1 \in E_1(\hat{r})$, and the verification that $f_2(x_1, g_f(x_1)) = g(f_1(x_1, gx_1))$ follows formally from the corresponding property for $\hat{f}$ and $\hat{g}$:

$$\begin{aligned} g(f_1(x_1, gx_1)) &= p_2 + \hat{g}(\hat{f}(x_1 - p_1, g(x_1) - p_2)) \\ &= p_2 + \hat{g}(\hat{f}_1(x_1 - p_1, \hat{g}(x_1 - p_1)) \\ &= p_2 + \hat{f}_2(x_1 - p_1, \hat{g}(x_1 - p_1)) \\ &= f_2(x_1, p_2 + \hat{g}(x_1 - p_1)) \\ &= f_2(x_1, g(x_1)). \end{aligned}$$

By 2.2, W_1 is the graph of $g = g_f$. The defining formula for g_f, and 2.3 for $\hat{f}$, prove 2.3 for f.

REMARK 1. The case $k = \infty$ follows from 2.3 because ε and δ are independent of k.

REMARK 2. The case $r = \infty$, that is, f is defined on all of E, presents difficulty in proving the global continuous dependence of the higher order derivatives of g_f on those of f. The trouble is that as $|x_1| \to \infty$, so does the number of iterates of f^{-1} needed to bring $(x_1, g_f(x_1))$ inside the small ball $E_1(r')$ where g_f is known to be C^k. Another proof, with a different induction hypothesis, can be devised.

REMARK 3. Suppose $f \in \mathcal{N}_\varepsilon^k(T)$, $\varepsilon < (1 - \tau)/(1 + \tau)$ and $f(0) = 0$. Put $\mathcal{G}_*^k = \{g \in \mathcal{G} \cap C^k : L(g) < 1\}$. Then $\Gamma_f : \mathcal{G}_*^k \to \mathcal{G}_*^k$ has g_f as attractive fixed point. Thus $\Gamma_f^n(g) \to g_f$, C^k, as $n \to \infty$, for any $g \in \mathcal{G}_*^k$. For $k = 1$ we proved this already: $\Delta\Gamma_f^n(g) = \Gamma_{\Delta f}^n(\Delta g) \to \Delta g_f$, C^0, and hence $\Gamma_f^n(g) \to g_f$, C^1. Assume the result holds for $k - 1 \geq 1$. Then for $0 < r' \leq r$ (as in the differentiability part of 2.3) we know $f' = f|E(r')$ has $\Delta f' \in \mathcal{N}_{\varepsilon'}^{k-1}(T \times T)$, $\varepsilon < \varepsilon' < (1 - \tau)/(1 + \tau)$ and $g' = g|E_1(r')$ has $\Delta g' \in \mathcal{G}_*^{k-1}(E \times E)$. Thus $\Gamma_{\Delta f'}^n(\Delta g') \to \Delta g_{f'}$, C^{k-1}, by induction. Hence $\Gamma_{f'}^n(g') \to g_{f'}$, C^k. But for $N \geq \log(r/r')/\log(\tau^{-1} - \varepsilon)$ we have, as in 2.3, graph $\Gamma_f^{n+N}(g) = f^N(\text{graph } \Gamma_{f'}^n(g')) \cap E(r)$ so that $\Gamma_f^{n+N}(g) \underset{n}{\to} f^N(\text{graph}(g_{f'})) \cap E(r) = \text{graph}(g_f)$, C^k.

We needed $L(g) < 1$ to insure $\Delta g' \in \mathcal{G}_*^{k-1}(E \times E)$ for some sufficiently small r'. If $L(g) = 1$ then $L(\Delta g')$ may be larger than 1 for all $r' > 0$. But of course the case of $L(g) = 1$ could be handled directly.

REMARK 4. If $f(0) \neq 0$ in the above, Γ_f does not map $\mathcal{G}$ into itself. If we enlarge $\mathcal{G}$ by letting $g(0) \neq 0$, with $|g(0)|$ appropriately small, we can prove directly that Γ_f is a well defined contraction and so has a fixed point, etc. The estimates are clearer, however, for $f(0) = 0$ and that is why we deduced the case $f(0) \neq 0$ formally from the case $f(0) = 0$ instead of proving both at once.

Now we indicate how to deduce the corresponding stable manifold theorem by inverting f.

2.4. STABLE MANIFOLD THEOREM FOR A POINT. *Let $0 < \tau < 1$ and $r > 0$ be given. There exist $\varepsilon > 0$ and $0 < \delta < \varepsilon$ with the following property. If $f : E(r) \to E$ satisfies*

$$L(f - T) < \varepsilon, \qquad |f(0)| < \delta,$$

for a hyperbolic isomorphism $T : E \to E$ of skewness τ then $W_2 = \bigcap_{n \geq 0} f^{-n}(E(r))$ is the graph of a unique function $g_{2f} : E_2(r) \to E_1(r)$. Moreover, $L(g_{2f}) \leq 1$ and g_{2f} is of class C^k if f is. The assignment $f \mapsto g_{2f}$ is continuous as a map $\mathcal{N}_\delta^k(T) \to C^k(E_2(r), E_1(r))$. The map $f|W_2 : W_2 \to W_2$ contracts W_2 into its interior.

PROOF. It suffices to show two things: 1^0. $L(f^{-1} - T^{-1}) \to 0$ and $f^{-1}(0) \to 0$ as

$L(f - T) \to 0$ and $f(0) \to 0$. 2°, $W_1(f^{-1}) = W_2(f)$. Since T^{-1} is linear it is justifiable to write

$$f^{-1} - T^{-1} = T^{-1}Tf^{-1} - T^{-1}ff^{-1} = T^{-1} \circ (T - f) \circ f^{-1},$$

so that

$$\begin{aligned} L(f^{-1} - T^{-1}) &\le \|T^{-1}\| L(T-f) L(f^{-1}) \\ &\le \|T^{-1}\| L(T-f)[\|T^{-1}\|^{-1} - L(T-f)]^{-1} \end{aligned}$$

by 1.4b. As $L(f - T) \to 0$ this tends to zero.

As for $f^{-1}(0)$ we have

$$\begin{aligned} |f^{-1}(0)| &\le |f^{-1}(0) - f^{-1}(p_f)| + |f^{-1}(p_f)| \\ &\le L(f^{-1})|p_f| |+p_f| \\ &\le ([\|T^{-1}\|^{-1} - L(f - T)]^{-1} + 1)\delta(1 - \tau - \varepsilon)^{-1} \to 0, \end{aligned}$$

as $\delta \to 0$ and $L(f - T) \to 0$.

Unfortunately, the second statement is not quite true, because f^{-1} is not defined on $E(r)$ but on $f(E(r))$.

Restricting to an $E(r') \subset fE(r)$ and arguing in a fashion similar to the $f(0) \neq 0$ part of 2.3 will show, for large enough N,

$$W_2(f) = f^{-N}(W_2(f|E(r')) = f^{-N}W_1(f^{-1}|E(r')).$$

The second equality is clear. Since $W_1(f^{-1}|E(r'))$ is well described by 2.3 we have proved 2.4.

3. **Stable manifolds for hyperbolic sets.** In this and the succeeding sections we adopt the following conventions.

M is a finite dimensional C^∞ Riemannian manifold; $U \subset M$ is an open set; $f: U \to M$ is a C^k embedding ($k \in Z_+$). We denote by $\Lambda \subset U$ a compact *invariant* set of f; that is, $f(\Lambda) = \Lambda$.

Notation. If $g: V \to V'$ is a C^1 map between smooth manifolds, we denote the *differential* of g by $Tg: TV \to TV'$. If $x \in V$, then V_x or T_xV denotes the tangent space to V at x, and $T_xg: T_xV \to T_{gx}V'$ is the restriction of Tg. For any subset $A \subset V$, $T_Ag: T_AV \to T_{gA}V'$ is the restriction of Tg to the tangent bundle of V over A.

If E is a Banach space and $g: V \to E$ a C^1 map, then $Dg_x: V_x \to E$ is the *derivative* of g at $x \in V$. This is by definition the composition

$$V_x \xrightarrow{T_xg} E_{gx} \xrightarrow{\xi_{fx}} E$$

where ξ_y denotes the canonical identification of the tangent space to E at $y \in E$ with E. $Dg: TV \to E$ is the map whose restriction to V_x is Dg_x.

If $p: X \to Y$ is a bundle, then $X_y = p^{-1}(y)$, the fiber over y. If $p: X \to Y$ and $p': X' \to Y'$ are bundles and $F: X \to X'$ takes fibers into fibers, covering $f: Y \to Y'$, then $F_y: X_y \to X'_{fy}$ is the restriction of F to X_y. We call such an F a *bundle map*.

If the vector bundles X and X' have Banach space structures on fibers, and $F: X \to X'$ is a linear bundle map, then we put $\|F\| = \sup_{y \in Y} \|F_y\|$.

We put $N_\alpha(x) = \{y | d(y, x) < \alpha\}$ in any metric space.

DEFINITION. The invariant set $\Lambda \subset U$ is *hyperbolic* for the map $f: U \to M$ if $T_\Lambda M$ has a splitting (Whitney sum decomposition) $T_\Lambda M = E^u \oplus E^s$ satisfying:

(1) E^u and E^s are invariant under Tf;

(2) there exist constants $c > 0$ and $0 < \tau < 1$ such that for all $n \in \mathbf{Z}_+$,

$$\max\{\|Tf^n|E^s\|, \|Tf^{-n}|E^u\|\} < c\tau^n.$$

We say Λ has *skewness* τ.

If the Riemann metric on M is such that in (2) we can take $c = 1$, then the metric is called *adapted* to Λ.

3.1. (MATHER, [10]). *If $\Lambda \subset U$ is a compact hyperbolic set for $M \supset U \xrightarrow{f} M$, then M has a smooth Riemann metric adapted to Λ.*

PROOF. Let $|v|$ denote the norm of $v \in TM$ in the metric for which Λ is hyperbolic. From (2) we have

$$|Tf^n(v)| \le c\tau^n|v| \quad \text{if } v \in E^s,$$
$$|Tf^n v| \ge c^{-1}\tau^{-n}|v| \quad \text{if } v \in E^u.$$

Let $q \in \mathbf{Z}_+$ be such that $c\tau^q < 1$. Define a new metric $\|v\|$ by

$$\|v\|^2 = \sum_{n=0}^{q-1} |Tf^n(v)|^2 \quad \text{if } v \in E^s$$
$$\|v\|^2 = \sum_{n=0}^{q-1} |Tf^{-n}(v)|^2 \quad \text{if } v \in E^u.$$

If $v \in E^s$, then $\|v\|^2 \le qc^2|v|^2$; and

$$\begin{aligned} \|Tf(v)\|^2 &= \sum_{n=1}^{q} |Tf^n(v)|^2 \\ &= \|v\|^2 - |v|^2 + |Tf^q(v)|^2 \\ &\le \|v\|^2 - [1 - (c\tau^q)^2]|v|^2 \\ &\le \|v\|^2(1 - [1 - c\tau^q]^2/qc^2). \end{aligned}$$

Thus $\|Tf(v)\| \le \sigma_0\|v\|$, where $\sigma_0^2 = 1 - [1 - c\tau^q]^2/qc^2$. Since $c\tau^q < 1$ we have $\sigma_0 < 1$. Similarly, replacing f by f^{-1}, we find that

$$\|Tf(v)\| \ge \sigma_0^{-1}\|v\| \quad \text{for } v \in E^u.$$

Now let $\sigma_0 < \sigma < 1$ and approximate $\|\cdot\|$ by a C^∞ metric $|||\cdot|||$ such that

$$|||Tf(v)||| \le \sigma|||v||| \quad \text{if } v \in E^s,$$

$$|||Tf(v)||| \ge \sigma^{-1} \; |||v||| \text{ if } v \in E^u.$$

The proof (due to Mather) of Theorem 3.1 is complete.

If $x \in U$ and $\beta > 0$ we put

$$\Sigma(x, \beta) = \bigcap_{n \ge 0} f^{-n}N_\beta(f^n x).$$

Thus $y \in \Sigma(x, \beta)$ if and only if $f^n(y)$ is defined, and $d(f^n y, f^n x) < \beta$, for all $n \geq 0$.

Let $\beta \geq \alpha > 0$. A submanifold $W \subset U$ is the *stable manifold through* $x \in U$ *of size* (β, α) provided $W = \Sigma(x, \beta) \cap N_\alpha(x)$. If $\beta = \alpha$ we say W has size β.

We shall show that for some $\beta \geq \alpha > 0$ every $x \in \Lambda$ has a stable manifold of size (β, α); if the metric is adapted to Λ, we can take $\beta = \alpha$.

In order to describe the sense in which the stable manifolds vary continuously, we make the following definition. A family $\{W_x\}_{x \in \Lambda}$ of C^k submanifolds of M is *continuous* if for each $x \in \Lambda$ there exists a neighborhood A of x in Λ and a (continuous) map $\phi : A \to C^k(D^n, M)$ such that ϕ_x maps D^n diffeomorphically onto a neighborhood of x in W_x, for each $x \in A$.

3.2. Stable Manifold Theorem For A Hyperbolic Set. *Let* $\Lambda \subset U$ *be a compact hyperbolic set for a* C^k *embedding* $M \supset U \xrightarrow{f} M$, $k \in Z_+$. *Then*

(a) *There exist numbers* $\beta \geq \alpha > 0$ *such that through each* $x \in \Lambda$ *there is a stable manifold* W_x *of size* (β, α).

(b) $\{W_x\}_{x \in \Lambda}$ *is a continuous family of* C^k *submanifolds.*

(c) *There exist numbers* $K > 0$ *and* $\lambda < 1$ *such that if* $y, z \in W_x$, *then* $d(f^n y, f^n z) \leq K\lambda^n d(y, z)$ *for all* $n \geq 0$.

(d) $W_x \cap W_y$ *is an open subset of* W_x *for all* $x, y \in \Lambda$.

(e) W_x *is tangent to* E^s_x *at* $x \in \Lambda$ (*where* $T_\Lambda M = E^u \oplus E^s$ *is the invariant splitting*).

(f) *If the metric on* M *is adapted to* Λ, *then* $\alpha = \beta$ *in* (a) *and* $K = 1$ *in* (c).

Proof. By Mather's theorem (3.1) we may assume the metric in M adapted to Λ. The basic idea behind the proof is to consider the Banach manifold $\mathcal{M}(\Lambda, M)$ of bounded maps $\Lambda \to M$. Define $\tilde{F} : \mathcal{M}(\Lambda, U) \to \mathcal{M}(\Lambda, M)$ by $\tilde{F}(h) = f \circ h \circ f^{-1}$. The inclusion $i : \Lambda \to M$ is a hyperbolic fixed point of $\tilde{F}$. If $\mathcal{W} \subset \mathcal{M}(\Lambda, U)$ is the stable manifold of $\tilde{F}$, put $\mathcal{W}_x = ev_x(\mathcal{W})$ = the set of points $h(x)$ for all $h \in \mathcal{W}$.

This definition is conceptually simple, but in order to prove that $\mathcal{W}_x$ is actually a submanifold it is more convenient to work in the exponential coordinate system of $\mathcal{M}$ at i.

Let $r > 0$ have the following property. For each $x \in \Lambda$, the exponential map $\exp_x : M_x(r) \to M$ is defined and maps $M_x(r)$ (= the ball of radius r in the tangent space M_x to M at x) diffeomorphically into $U \cap f^{-1}U$. This is possible since Λ is compact.

Let $V_r = \{y \in T_\Lambda M \mid |y| \leq r\}$. Define $\tilde{F} : V_r \to T_\Lambda M$ by

$$\tilde{F} | V_r \cap M_x = \exp_{fx}^{-1} \circ f \circ \exp_x.$$

The following diagram commutes:

$$\begin{array}{ccccc} & V_r & \xrightarrow{\tilde{F}} & T_\Lambda M & \\ (p, \exp) & \downarrow & & \downarrow & (p, \exp) \\ & \Lambda \times f^{-1}U & \xrightarrow{f \times f} & \Lambda \times U & \\ & \downarrow & & \downarrow & \\ & \Lambda & \xrightarrow{f} & \Lambda & \end{array}$$

where $p : T_\Lambda M \to \Lambda$ is the bundle projection.

Let S be the Banach space of all bounded, possibly discontinuous sections of $T_\Lambda M$. Let $S(r)$ be the (closed) ball of radius r in S around 0. Then F induces a map $F:S(r) \to S$ by $F(\sigma) = \tilde{F} \circ \sigma \circ f^{-1}$. More explicitly, if $x \in \Lambda$ and $\sigma \in S(r)$ then $F(\sigma)x = (\exp_x)^{-1} f \exp_{f^{-1}x}\sigma(f^{-1}x)$.

It is left to the reader to verify that F is C^k. The derivative of F at $0 \in S(r)$ is the linear map $DF_0:S \to S$, defined by the formula $DF_0(\sigma) = (Tf) \circ \sigma \circ (f^{-1}|\Lambda)$.

It is easy to see that F is hyperbolic. The splitting of S is $S^u \oplus S^s$ where

$$S^u = \{\sigma \in S | \sigma(\Lambda) \subset E^u\}, \text{and } S^s = \{\sigma \in S | \sigma(\Lambda) \subset E^s\}.$$

Therefore by Theorem 2.3, F has a stable manifold function $G:S^s(r) \to S^u(r)$. (It may be necessary to replace r by a smaller number. We assume this done, and remark that the smallness of r depends only on the constants of hyperbolicity, C and τ, and on the first order "Taylor expansions" of f at points $x \in \Lambda$: if $v \in M_x$, put $f(\exp v) = \exp_{fx}((Tf)v + o(v))$. Hence a single r can be chosen for a whole neighborhood of f in $C^1(U, M)$.)

We recall from 2.2 and 2.4 that the stable manifold function G is characterized as follows: Given $\sigma \in S^s(r)$, $G(\sigma)$ is the unique section $\tau \in S^u(r)$ such that $F^n(\tau, \sigma)$ is defined and lies in $S(r)$ for all $n \geq 0$.

LEMMA. *There is a unique map $H:E^s(r) \to E^u(r)$ covering 1_Λ such that $G(\sigma) = H \circ \sigma$ for all $\sigma \in S^s(r)$.*

PROOF. Given $y \in T_\Lambda M$, define $\sigma_y \in S$ by

$$\sigma_y(x) = 0 \quad \text{if } x \neq p(y),$$
$$\quad\;\; = y \quad \text{if } x = p(y).$$

For $y \in E_x^s(r)$ put $H(y) = G(\sigma_y)x$. Now suppose $\sigma(x) = y$. Then

$$|\tilde{F}^n(G(\sigma)x, y)| \leq r \quad \text{for all } n,$$

and also

$$|\tilde{F}^n(H(y), y)| \leq r \quad \text{for all } n.$$

From the characterization of $G(\sigma)$ given above, we must have $G(\sigma)x = H(y)$.

LEMMA. *H is continuous, and C^k on each fiber $E_x^s(r)$.*

PROOF. Let $\Sigma(y) = \sigma_y$. Then H is the composition of the continuous maps:

$$H:E^s(r) \xrightarrow{(\Sigma, p)} S^s(r) \times \Lambda \xrightarrow{(G, 1)} S^u(r) \times \Lambda \xrightarrow{ev} E^u(r).$$

Since $\Sigma:E^s \to S^s$ and $ev_x:S^u \to E^u$ are linear and G is C^k, it follows that H is C^k on fibers.

Now let $\mathscr{W}$ be the graph of H, that is,

$$\mathscr{W} = \{(Hy, y) \in E^u(r) \times E^s(r) | y \in E^s(r)\}.$$

For each $x \in \Lambda$ put $\mathscr{W}_x = \mathscr{W} \cap M_x(r)$. Then $\mathscr{W}_x$ is a C^k submanifold of $M_x(r)$. Therefore $W_x = \exp_x(\mathscr{W}_x - \partial \mathscr{W}_x)$ is a C^k submanifold of U. It follows from the composition formula for H given above that $\{W_x\}_{x \in \Lambda}$ is a continuous family of C^k submanifolds.

To see that W_x is a stable manifold for x, choose $\beta > 0$ so small that $N_\beta(z) \subset \exp_z(M_r(z))$ for all $z \in \Lambda$. If $y \in \Sigma(x, \beta)$, that is, $f^n(y)$ is defined and in $N_\beta(f^n(x))$ for all $n \geq 0$, then define a section σ of $T_\Lambda M(r)$ by

$$\begin{aligned} \sigma(z) &= \exp_x^{-1}(y) \quad \text{if } z = x \\ &= 0 \qquad\qquad\;\, \text{if } z \neq x. \end{aligned}$$

Then $F^n(\sigma)$ is defined and in $S(r)$ for all $n \geq 0$. Therefore σ is in the stable manifold of G. This means that $\exp_x^{-1} y = H(\sigma)x$ which in turn means that $y \in W_x$.

Part (c) of Theorem 3.2 follows from the analogous fact for stable manifolds of hyperbolic fixed points; (d), (e) and (f) are left to the reader.[2]

4. **Criteria for hyperbolicity.** In this section, which is basically independent of the preceding ones, we establish tests for a linear map or an invariant set to be hyperbolic. These are based on criteria for a linear operator on a Banach space to be hyperbolic. The *universality* of these criteria is important: they are valid in every Banach space, and no special properties of the Banach space are used. The basic estimate is 4.7; the criteria are found in 4.8–4.10.

NOTATION. If E and F are Banach spaces, $L(E, F)$ is the Banach space of linear maps $E \to F$; its unit ball is $L_1(E, F)$. $\text{Inv}(E, F)$ is the open subset of invertible linear maps.

The following universal estimate is well known.

4.1. LEMMA. *Let $P \in L(E, E)$ have norm < 1. Then $I + P: E \to E$ is invertible and* $\|(I + P)^{-1}\| \leq (1 - \|P\|)^{-1}$.

Next we state an exercise in differential calculus on noncommutative Banach algebras.

4.2. LEMMA.

(a) *Let B be a ball in a Banach space V. Let $f: B \to L(E, F)$ and $g: B \to L(F, G)$ be differentiable at $b \in B$. The map $h: B \to L(E, G)$, defined by $h(x) = g(x) \circ f(x)$ is also differentiable at b, and $Dh_b \in L(V, L(E, G))$ is the map assigning to $x \in V$ the linear map*

$$g(b) \circ Df_b(x) + Dg_b(x) \circ f(b): E \to G.$$

(b) *Define $\iota: \text{inv}(E, F) \to \text{inv}(F, E)$ by $\iota(T) = T^{-1}$. Then ι is differentiable and if $T \in \text{inv}(E, F)$ then $D\iota_T: L(E, F) \to L(E, F)$ is the linear map $S \mapsto -T^{-1}ST^{-1}$.*

PROOF Left to reader. See [4].

4.3. LEMMA. *Let E_i and F_i be Banach spaces, $i = 1, 2$. Let $T: E_1 \times E_2 \to F_1 \times F_2$ be defined by the matrix of linear maps*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

so that $T(x_1, x_2) = (Ax_1 + Bx_2, Cx_1 + Dx_2)$. *Suppose* $A: E_1 \to F_1$ *is invertible. Put* $u = \|B\| \cdot \|A^{-1}\|$, *and suppose* $u < 1$. *Then the graph transform*

$$\Gamma_T : L_1(E_1, E_2) \to L(F_1, F_2),$$

is well defined by the formula

$$\Gamma_T(P) = (C + DP)(A + BP)^{-1}.$$

Moreover the Lipschitz constant $L(\Gamma_T)$ *satisfies*

$$L(\Gamma_T) \le (\|CA^{-1}\| + \|D\| \cdot \|A^{-1}\|)u(1 - u)^{-2} + \|D\| \cdot \|A^{-1}\|(1 - u)^{-1}.$$

PROOF. If $P \in L_1(E_1, E_2)$ then $\|BPA^{-1}\| \le \|B\| \cdot \|A^{-1}\| = u < 1$. Therefore $I + BPA^{-1}$ is invertible and

$$\|(I+BPA^{-1})^{-1}\| \le (1 - u)^{-1} \tag{1}$$

by Lemma 4.1.

Thus $(A + BP)^{-1} = A^{-1}(I + BPA^{-1})^{-1}$ exists, so that Γ_T is well defined. The derivative $(D\Gamma_T)_P : L(E_1, E_2) \to L(F_1, F_2)$ of Γ_T at P takes $X \in L(E_1, E_2)$ into

$$-(C + DP)(A + BP)^{-1}BX(A + BP)^{-1} + DX(A + BP)^{-1}, \tag{2}$$

by Lemma 4.2. Put $(A + BP)^{-1} = A^{-1}Q$, $Q = (I + BPA^{-1})^{-1}$. Then $\|Q\| \le (1 - u)^{-1}$ by (1), and from (2) we have

$$(D\Gamma_T)_P(X) = (CA^{-1} + DPA^{-1})QBXA^{-1}Q + DXA^{-1}Q. \tag{3}$$

Since $\|P\| \le 1$ and $\|Q\| \le (1 - u)^{-1}$, the result follows.

For convenience, put

$$\|CA^{-1}\| = w, \quad \|D\| \cdot \|A^{-1}\| = v, \quad \|B\| \cdot \|A^{-1}\| = u.$$

Then Lemma 4.1 implies

$$L(\Gamma_T) \le (w + v)u(1 - u)^{-2} + v(1 - u)^{-1}. \tag{4}$$

4.4. LEMMA. *If* $u < 1$ *and* $P \in L_1(E_1, F_1)$ *then* $\|\Gamma_T(P)\| \le (w + v)(1 - u)^{-1}$.

PROOF.

$$\begin{aligned}\|\Gamma_T(P)\| &= \|(CA^{-1} + DPA^{-1})(I + BPA^{-1})^{-1}\| \\ &\le (\|CA^{-1}\| + \|D\| \cdot \|A^{-1}\|) \cdot \|Q\| \\ &\le (w + v)(1 - u)^{-1}.\end{aligned}$$

4.5. PROPOSITION. *With the above notation, let* $F_i = E_i$ $(i = 1, 2)$. *Suppose*

$$2u < 1 - v, \tag{a}$$

and

$$u + v + w < 1. \tag{b}$$

Then $\Gamma_T : L_1(E_1, E_2) \to L_1(E_1, E_2)$ *is a well-defined contraction of Lipschitz constant* $\le (u + v)(1 - u)^{-1} < 1$. *Consequently* Γ_T *has a unique fixed point in* $L_1(E_1, E_2)$, *which depends continuously on* T.

PROOF. By Lemma 4.3, Γ_T is well defined on $L_1(E_1, E_2)$; by (b) and Lemma 4.4, the image of Γ_T lies in $L_1(E_1, E_2)$. Combining (a), (b) and 4.4. yields the estimate of $L(\Gamma_T)$.

Now let $W \subset E_1 \times E_2$ be the graph of the fixed point $G \in L_1(E_1, E_2)$ of Γ_T. We give W the Banach norm it inherits as a closed subspace of $E_1 \times E_2$.

4.6. LEMMA. *Assume* $\|A^{-1}\|(1-u)^{-1} < 1$. *Then* $T|W$ *is expanding; in fact*

$$|Ty| \geq (1-u)\|A^{-1}\|^{-1}|y| \quad \text{if } y \in W.$$

PROOF. Put $y = (x, Gx)$ with $x \in E_1$. Then $|y| = \max(|x|, |Gx|) = |x|$ since $\|G\| \leq 1$. Similarly

$$\begin{aligned} |Ty| &= |(Ax + BGx, Cx + DGx)| \\ &= |Ax + BGx| \\ &\geq \|(A + BG)^{-1}\|^{-1}|x| \\ &\geq \|A^{-1}\|^{-1} \cdot \|(I + BGA^{-1})^{-1}\|^{-1} \cdot |x|. \end{aligned}$$

The lemma follows since

$$\|(I + BGA^{-1})^{-1}\| \leq (1-u)^{-1}.$$

We summarize these facts:

4.7. PROPOSITION. *Let*

$$T = \begin{bmatrix} A & B \\ C & D \end{bmatrix} : E_1 \times E_2 \to E_1 \times E_2$$

be as above. Suppose

(a) $\|A^{-1}\| < 1$.
(b) $u + v + w < 1$.
(c) $2u < 1 - v$,

where $w = \|CA^{-1}\|$, $v = \|D\| \cdot \|A^{-1}\|$ *and* $u = \|B\| \cdot \|A^{-1}\|$. *Then the graph transform*

$$\Gamma_T : L_1(E_1 \times E_2) \to L_1(E_1 \times E_2),$$

is a well-defined contraction. If $W \subset E_1 \times E_2$ *is the graph of the fixed point* G_T *of* Γ_T, *then* $T|W$ *is an expansion. Moreover* G_T *depends continuously on* T, *and* $\|G_T\| \leq (w + v)/(1-u)$.

We now derive a perturbation criterion for hyperbolicity.

4.8. THEOREM. *Given* $0 < \tau < 1$, *there exists* $\varepsilon > 0$ *with the following property. Let* E_1 *and* E_2 *be Banach spaces,*

$$T = \begin{bmatrix} A & B \\ C & D \end{bmatrix} : E_1 \times E_2 \to E_1 \times E_2$$

a linear map with $A : E_1 \to E_1$ *and* $D : E_2 \to E_2$ *invertible. Suppose*

$$\max\{\|A^{-1}\|, \|D\|\} < \tau + \varepsilon \text{ and } \max\{\|B\|, \|C\|\} < \varepsilon.$$

Then T is hyperbolic (for some splitting of $E_1 \times E_2$*).*

PROOF. By choosing ε small enough we may assume T invertible, and apply Proposition 4.7 to both T and T^{-1} to get the expanding and contracting invariant subspaces for T.

Another universal criterion is the following.

4.9. PROPOSITION. *Let* $0 < \tau < 1$ *and* $\varepsilon > 0$. *There exists* $\delta > 0$ *with the following property. Let* E_i *and* F_i *be Banach spaces* $(i = 1, 2)$, *and* $T_i: E_i \to F_i$ *invertible linear maps such that*

(a) $\max\{\|T_1^{-1}\|, \|T_2\|\} \leq \tau < 1$.

Let

$$H = \begin{bmatrix} P & Q \\ R & S \end{bmatrix} : F_1 \times F_2 \to E_1 \times E_2$$

be a linear map with $P: F_1 \to E_1$ *invertible, satisfying*

(b) $\max\{\|P^{-1}\|^{-1} - 1, \|Q\|, \|R\|, \|S\| - 1\} < \delta$.

Then the map $HT: E_1 \times E_2 \to E_1 \times E_2$ *is hyperbolic for some splitting* $E^u \times E^s$ *of* $E_1 \times E_2$. *Moreover if* $G^u: E_1 \to E_2$ *and* $G^s: E_2 \to E_1$ *are the unstable and stable manifold functions of* HT, *then*

(c) $\max\{\|G^u\|, \|G^s\|\} < \varepsilon$.

(d) *Analogous statements hold for*

$$TH: F_1 \times F_2 \to F_1 \times F_2.$$

PROOF. Let

$$HT = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} PT_1 & QT_2 \\ RT_1 & ST_2 \end{bmatrix}.$$

By Proposition 4.6, HT has an unstable manifold provided

$$\|A^{-1}\| < 1,$$

$$\|A^{-1}\| \cdot \|D\| + \|A^{-1}\| \cdot \|B\| + \|CA^{-1}\| < 1,$$

and

$$2\|A^{-1}\| \cdot \|B\| < 1 - \|A^{-1}\| \cdot \|D\|.$$

This will be true provided

(5) $$\tau\|P\| < 1,$$

(6) $$\|S\|\tau + \|Q\|\tau + \|R\| \cdot \|P^{-1}\| < 1,$$

and

(7) $$2\|Q\|\tau < 1 - \|S\|\tau.$$

If δ is sufficiently small then (b) implies (5), (6) and (7). Moreover

$$\|G^u\| \leq (w + v)(1 - u)^{-1} \leq (\|R\| \cdot \|P^{-1}\| + \|Q\|\tau)(1 - \|S\|\tau)^{-1},$$

which is $<\varepsilon$ if δ is sufficiently small.

Applying this result to $H^{-1}T^{-1}$ shows that if δ is sufficiently small then $TH: F_1 \times F_2 \to F_1 \times F_2$ has a stable manifold, and a stable manifold function of norm $< \varepsilon$. Similar reasoning shows that TH has an unstable manifold and HT a stable manifold, etc.

REMARK. Notice that the estimate of $\|CA^{-1}\|$ was used (from Proposition 4.7) rather than an estimate of $\|C\| \cdot \|A^{-1}\| = \|RT_1\| \cdot \|T_1^{-1}P^{-1}\|$. The latter requires a knowledge of $\|T_1\|$ whereas the former does not.

Next we apply Theorem 4.8 to get a criterion for a hyperbolic set.

4.10. PROPOSITION *Let $M \supset V \xrightarrow{g} M$ be a C^1 embedding of an open set V. Let $X \subset V$ be invariant under g and g^{-1}. Let $E_1 \oplus E_2$ be a splitting of $T_X M$. Put*

$$T_X g = \begin{bmatrix} A & B \\ C & D \end{bmatrix} : E_1 \oplus E_2 \to E_1 \oplus E_2,$$

where A, B, C, D are bundle maps covering g. If there exist $0 < \tau < 1$ and $\varepsilon > 0$ satisfying Theorem 4.8, and also

(a) $\max\{\|A^{-1}\|, \|D\|\} < \tau + \varepsilon$,

(b) $\max\{\|B\|, \|D\|\} < \varepsilon$,

then X is a hyperbolic set.

PROOF. Let $C_0(T_X M) = C_0$ be the Banach space of bounded continuous sections of $T_X M$. Let $F: C_0 \to C_0$ be induced by g. That is, $F(\sigma) = T_X g \circ \sigma \circ g^{-1}$. If we write $C_0 = C_0(E_1|X) \times C_0(E_2|X)$, then Proposition 4.8 shows that F is hyperbolic for some splitting of C_0. An obvious extension of a theorem of J. Mather [12, Appendix] concerning Anosov diffeomorphisms shows that this suffices for X to be a hyperbolic set. Alternatively, the proof of the existence theorem (3.2) for stable manifolds could be imitated. (This idea is due to S. Smale; in fact it suggested the proof of 3.1 given here.)

5. **Hyperbolicity of submanifolds.** Let $\Lambda \subset U$ be a hyperbolic set for $M \supset U \xrightarrow{f} M$, and $V \subset \Lambda$ a smooth submanifold invariant under f and f^{-1}. No examples are known for which $f|V: V \to V$ is not Anosov; on the other hand there is no proof that $f|V$ must be Anosov (which means that V is a hyperbolic set for $f|V$). We prove a partial result in this direction.

5.1. THEOREM. *Let $p: E \to V$ be a finite dimensional vector bundle. Let $T: E \to E$ be a linear bundle automorphism covering $f: V \to V$. Let $F \subset E$ be an invariant subbundle over V. Let $\Omega \subset V$ be the set of nonwandering points. If T is hyperbolic, then $T|F_\Omega$ is hyperbolic.*

PROOF. Let $E^u \oplus E^s$ be the hyperbolic splitting of E. Let the metric on E be adapted to T: there exists $0 < \tau < 1$ such that $|Tv| \leq \tau|v|$ if $v \in E^s$, $|Tv| \geq \tau^{-1}|v|$ if $v \in E^u$. (See 3.1.)

We shall prove that if $x \in \Omega$, then $F_x = (F_x \cap E_x^u) \times (F_x \cap E_x^s)$; this suffices to prove $T|F_\Omega$ hyperbolic. We do this by proving

$$\text{(1)} \qquad \dim E_x^u \cap F_x + \dim E_x^s \cap F_x \geq \dim F_x.$$

Let $x \in \Omega$, and let $W \subset V$ be a set over which the bundle pair (E, F) is trivial. Let $\phi:(E_W, F_W) \to (R^n, R^m)$ be a trivialization. For each $y \in W$, $\phi_y = \phi|E_y$ maps E_y isomorphically onto R^n, and F_y onto R^m.

Suppose $y \in W$ and $z = f^k(y) \in W$. Then

$$T_{k,y} = \phi_y^{-1}\phi_z(F^k)_y:(E_y, F_y) \to (E_y, F_y),$$

is a linear automorphism. In terms of the splitting $E_y = E_y^u \times E_y^s$, $T_{k,y}$ is represented by a matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

We want to be in the situation of Proposition 4.9, with T and H of 4.9 represented by $(F^k)_y$ and $\phi_y^{-1}\phi_z$ respectively.

Since x is nonwandering, we can find $y \in U$ and $k \in Z_+$ such that y and $z = f^k(y)$ are as close to x as desired. In particular we can make $\phi_y^{-1}\phi_z$ as close to an isometry as desired, so that (b) of 4.9 will hold; we can simultaneously take k as large as necessary for 4.9(a) to hold. It follows that $T_{k,y}:E_y \to E_y$ will be hyperbolic. Since $F_y \subset E_y$ is invariant under $T_{k,y}$, *and* E_y *is finite dimensional,* a simple eigenvalue argument shows that $T_{k,y}|F_y$ is hyperbolic. Moreover the stable and unstable manifolds of $T_{k,y}|F_y$ must be the intersection of F_y with the stable and unstable manifolds B_y^s, B_y^u of $T_{k,y}$. And (c) of 4.9 shows that as y and z approach x, B_y^s and B_y^u approach E_x^s and E_x^u. Since $\dim B_y^u \cap F_y + \dim B_y^s \cap F_y = \dim F_y$, (1) follows and the theorem is proved.

5.2. Corollary. *Let $V \subset M$ be a compact invariant C^1 submanifold of an Anosov diffeomorphism $f:M \to M$. Then the nonwandering set Ω_0 of $f|V$ is hyperbolic.*

Remark 1. If $F \subset E$ is a subspace invariant under T and T^{-1}, then every eigenvalue of T is clearly an eigenvalue of $T|F$. Therefore $T|F$ is hyperbolic if T is hyperbolic and spectrum (T) consists entirely of eigenvalues as in the finite dimensional case.

Remark 2. The following infinite dimensional example, due to W. Badé, shows that $T|F$ may fail to be hyperbolic.

Let C be the complex field. Let

$$A = \{z \in C \mid \tfrac{1}{4} \leq |z| \leq \tfrac{1}{2}\}, \qquad B = \{z \in C \mid 2 \leq |z| \leq 3\}.$$

Let E be the Hilbert space of complex functions which are continuous on $A \cup B$ and analytic on $\operatorname{int}(A \cup B)$. Define $T:E \to E$ by $T(f)z = zf(z)$. Then T is hyperbolic; the invariant splitting is found by setting

$$E^u = \{f \in E \mid f(A) = 0\}, \qquad E^s = \{f \in E \mid f(B) = 0\}.$$

Let $F \subset E$ be the subspace comprising those functions that extend to a function analytic on $\{z | \frac{1}{4} < |z| < 3\}$.

Then $TF = F = T^{-1}F$. However, the constant function 1 belongs to F, but is not in the image of $(T - I)|F: F \to F$. Hence the complex number 1 is in spectrum $(T|F)$.

6. **Smoothness of splittings.** In order to study the smoothness of the splitting $T_\Lambda M = E^u \oplus E^s$, or more precisely, the smoothness of the functions assigning to each $x \in \Lambda$ the subspaces E_x^u and E_x^s, we must study the smoothness of certain sections of vector bundles. There is no "natural" metric on a vector bundle $p: E \to X$ in which to express Hölder conditions, but there is a natural class of metrics which we now define.

DEFINITION. Let $p: E \to X$ be a vector bundle over a metric space X. A metric d on E is *admissible* if there is a complementary bundle E' over X, and an isomorphism $h: E \oplus E' \to X \times A$ to a product vector bundle, where A is a Banach space, such that d is induced from the product metric on $X \times A$.

6.1. THEOREM. *Let $p: Y \to X$ be a vector bundle over a metric space X endowed with an admissible metric. Let $D \subset Y$ be the unit ball bundle, and $F: D \to D$ a map covering a homeomorphism $f: X \to X$. Suppose $0 \le \kappa < 1$ and that for each $x \in X$, the restriction $F_x: D_x \to D_{fx}$ has Lipschitz constant $\le \kappa$. Then*

(a) *There is a unique section $g_0: X \to D_0$ whose image is invariant under F.*

(b) *Let $L(f^{-1}) = \lambda < \infty$; let $0 < \alpha \le 1$ be such that $\kappa\lambda^\alpha < 1$. Then g_0 satisfies a Holder condition of exponent α.*

(c) *Suppose X is a smooth manifold, E is a smooth vector bundle, and F, f are C^1. If $\kappa\lambda < 1$ then g_0 is C^1.*

PROOF. Let $\mathscr{G}$ be the unit ball in the Banach space of bounded continuous sections of Y. Define $\Phi: \mathscr{G} \to \mathscr{G}$ by $\Phi(g) = Fgf^{-1}$. Then $L(\Phi) \le \kappa < 1$; hence Φ has a unique fixed point g_0. This proves (a).

Before proving (b) and (c) we remark that Y may be assumed trivial. For let Z be a bundle over X such that $Y \oplus Z$ is trivial, and define F' to be the composition

$$F': Y \oplus Z \xrightarrow{\pi} Y \xrightarrow{F} Y \xrightarrow{i} Y \oplus Z$$

where π is the projection and i the inclusion. Then F' satisfies the same hypotheses as F, and the unique invariant section of F is ig_0. Henceforth, we assume $Y = X \times E$ where E is a Banach space. We write $F(x, y) = (fx, f_x y)$.

The proof of (c) is more intuitive than that of (b), so we do it first. Assume X is a Riemannian manifold; let $B \subset E$ be the unit ball. Sections are now maps; $g_0: X \to B$ is the unique map whose graph is invariant under $F: X \times B \to X \times B$.

Let $\mathscr{H}$ be the Banach space of linear bundle maps $H: TX \to X \times E$ covering 1_X and having finite norm $\|H\| = \sup_{x \in X} \|H_x\|$, where $H_x: T_xX \to E$ is defined by setting $H(v) = (x, H_x v)$ for $v \in T_xX$.

For each $g \in \mathscr{G}$ define a linear map $\psi_g: \mathscr{H} \to \mathscr{H}$ as follows. If $v \in T_xX$ put $f^{-1}(x) = y \in X$ and $(Tf^{-1})v = w \in T_yX$. If $H \in \mathscr{H}$ define

$$(\Psi_g H)_x v = D(\pi_2 \circ F)_{(y, gy)}(w, H_y w).$$

Here $\pi_2 : X \times E \to E$ is the projection; $D(\pi_2 \circ F)_{(y,gy)}$ is the linear map from the tangent space at $X \times E$ at (y, gy) to E that is the derivative of $\pi_2 \circ F$; and $(w, H_y w)$ is a tangent vector to $X \times E$ at (y, gy). Observe that if $g \in \mathscr{G}$ is C^1, then $\Psi_g(Dg) = D(\Phi g)$.

We claim $L(\Psi_g) \leq \lambda\kappa$. By definition if $H, K \in \mathscr{H}$ then $\|H - K\| = \sup_{x \in X_0} \|H_x - K_x\|$. Observe that

$$(\psi_g H - \psi_g K)_x v = (DF_y)_{gy}(H_y - K_y)(T_x f^{-1})v.$$

Since $\|(DF_y)_{gy}\| \leq \kappa$ if $L(F_y) \leq \kappa$, and $\|T_x f^{-1}\| \leq L(f^{-1}) \leq \lambda$, we have

$$\|\Psi_g H - \Psi_g K\| \leq \kappa\lambda \|H - K\|.$$

Therefore Ψ_g is a contraction if $\kappa\lambda < 1$. By the Fiber Contraction Theorem (1.2) the map $\Psi : \mathscr{G} \times \mathscr{H} \to \mathscr{G} \times \mathscr{H}$, defined by $(g, H) \mapsto (\Phi g, \Psi_g H)$, has an attractive fixed point (g_0, H_0). If $g \in \mathscr{G}$ is C^1, then $\Psi^n(g, Dg) = (\Phi^n g, D\Phi^n g)$. Therefore $D(\Phi^n g)$ converges, and so g_0 is C^1. This proves 6.1(c).

To prove (b) we assume F is defined on all of $X \times E$. To see that there is no loss of generality, let $r : B \to E$ be the radial retraction

$$\begin{aligned} r(x) &= x \quad \text{if } |x| \leq 1 \\ &= x/|x| \text{ if } |x| \geq 1. \end{aligned}$$

Then $L(r) \leq 2$, and $L((Fr)_x^n) = L((F^n)_x r) = 2\kappa^n$, which is <1 if n is large enough. If $\kappa\lambda^\alpha < 1$, then $(2\kappa^n)(\lambda^n)^\alpha = 2(\kappa\lambda^\alpha)^n$ is also < 1 if n is large enough. Therefore f and F may be replaced by f^n and $F^n r$; g_0 and α stay the same, and also $L((F^n r)_x) L(f^{-n})^\alpha < 1$. Therefore we assume $F : X \times E \to X \times B$ given covering f, with $L(F_x) \leq \kappa < 1$; and $L(f^{-1}) \leq \lambda$ with $\kappa\lambda^\alpha < 1$.

In order to imitate the proof of (c), we replace Dg by $\Delta g : X \times X \to E$, defined by $\Delta g(x, y) = g(x) - g(y)$. We proceed as follows. Let $\mathscr{G}$ and $\Phi : \mathscr{G} \to \mathscr{G}$ be as before. Let $\mathscr{H}$ be the Banach space of bounded continuous maps $H : X \times X \to E$, such that $H(x, x) = 0$ for all x and the following norm is finite:

$$\|H\| = \sup_{x \neq y} |H(x, y)|/d(x, y)^\alpha + \sup_{(x,y)} |H(x, y)|.$$

The natural map from $\mathscr{H}$ to the Banach space of bounded continuous maps $X \times X \to E$ is continuous and takes closed bounded sets onto closed bounded sets. If $H \in \mathscr{H}$ and $x \in X$ define $H_x : X \to E$ by $H_x(y) = H(x, y)$.

Given $g \in \mathscr{G}$, define $\Psi_g : \mathscr{H} \to \mathscr{H}$ by

$$(\Psi_g H)_x = \Phi(g + H_{f^{-1}x}) - \Phi(g).$$

If $g \in \mathscr{G}$ satisfies an α-Hölder condition, define $\Delta g \in \mathscr{G}$ by $\Delta g(x, y) = g(x) - g(y)$. Observe that $\Psi_g(\Delta g) = \Delta(\Phi g)$. Define $\Psi : \mathscr{G} \times \mathscr{H} \to \mathscr{G} \times \mathscr{H}$ by $\Psi(g, H) = (\Phi g, \Psi_g H)$.

We show now that $L(\Psi_g) \leq \kappa\lambda^\alpha$:

$$\begin{aligned} d(x, y)^\alpha \|\Psi_g(H) - \Psi_g(K)\| &= \sup_{x \neq y} |(\Psi_g H)_x y - (\Psi_g K)_x y| \\ &= \sup_{x \neq y} |F_{f^{-1}x}(g + H_{f^{-1}x})y - F_{f^{-1}x}(g + K_{f^{-1}x})y| \end{aligned}$$

$$\leq \kappa \sup_{x \neq y} |H_{f^{-1}x}(f^{-1}y) - K_{f^{-1}x}(f^{-1}y)|$$

$$\leq \kappa \|H - K\| [d(f^{-1}y, f^{-1}x)]^{\alpha}$$

$$\leq \kappa \|H - K\| \lambda^{\alpha} d(x, y)^{\alpha}.$$

Therefore $\|\Psi_g(H) - \Psi_g(K)\| \leq \kappa\lambda^{\alpha}\|H - K\|$.

Since $\kappa\lambda^{\alpha} < 1$ it follows from the Fiber Contraction Theorem (1.2) and the completeness of $\mathscr{H}$ that $\Psi: \mathscr{G} \times \mathscr{H} \to \mathscr{G} \times \mathscr{H}$ has a unique attractive fixed point. If $g \in \mathscr{G}$ is α-Hölder, then $\Phi^n g \to g_0$, while $\Delta(\Phi^n g)$ converges in $\mathscr{H}$, as $n \to \infty$. But therefore $\Delta(\Phi^n g)$ converges in the uniform topology, so it must converge to $\Delta(g_0)$. Therefore g_0 is α-Hölder. Q.E.D.

It is useful to generalize 6.1 to the case where f maps a subspace X_0 of X homeomorphically onto X. For simplicity we deal only with the smooth case.

6.2. THEOREM. *Let $p: E \to X$ be a C^1 vector bundle over a C^1 manifold X. Let $X_0 \subset X$ be an open set and $f: X_0 \to X$ a C^1 diffeomorphism. Let $D \subset E$ be the unit ball bundle in an admissible metric and put $D_0 = p^{-1}X_0$. Suppose $F: D_0 \to \operatorname{int} D_0$ is a C^1 map covering f such that $L(F_x) \leq \kappa < 1$ for each $x \in X_0$. Then*

(a) *there exists a unique section g_0 of D_0 that is invariant under F in the sense that $g_0(X_0) \subset Fg_0(X_0)$.*

(b) *Let $L(f^{-1}) = \lambda < \infty$. If $\kappa\lambda < 1$ then g_0 is C^1.*

PROOF. The proof is practically identical to that of 6.1(c) and is left to the reader.

Now let $\Lambda \subset U$ be a compact hyperbolic set for $M \supset U \xrightarrow{f} M$. Let $T_\Lambda M = E^s \oplus E^u$ be the invariant splitting, with E^s contracting and E^u expanding. We define four quantities:

$$a = \|Tf^{-1}|E^u\| < 1 \qquad b = \|Tf|E^s\| < 1$$
$$c = \|Tf|E^u\| > 1 \qquad d = \|Tf^{-1}|E^s\| > 1.$$

If $\Lambda = U = M$ then f is called an *Anosov* diffeomorphism of M. In this case the stable and unstable manifolds give two topological foliations of M. According to [3] these are not always C^1, although they are "absolutely continuous."

6.3. THEOREM. *Let $f: M \to M$ be a C^2 Anosov diffeomorphism. Then*

(a) *The stable foliation is α-Hölder where $0 < \alpha \leq 1$ and $abc^{\alpha} < 1$; the unstable foliation is β-Hölder where $0 < \beta \leq 1$ and $abd^{\beta} < 1$.*

(b) *If $abc < 1$ then the stable foliation is C^1. In particular, the stable foliation is C^1 in these two cases:* (i) *the stable manifolds have codimension one;* (ii) $\dim M = 3$ *and f preserves the Riemannian measure in M; in this case the unstable foliation is also C^1.*

PROOF. Give TM a C^1 splitting $F^s \oplus F^u$ approximating $E^s \oplus E^u$. For each $x \in M$ put $L_x = L(F^s, F^u)$; then E_x^s is the graph of an element $\lambda_x \in L_x$. Define $\Gamma_x: L_x(1) \to L_y(1)$, $y = f^{-1}x$, to be the graph transform induced by

$$T_xf^{-1} = \begin{bmatrix} A_x & B_x \\ C_x & D_x \end{bmatrix} : F^s_x \times F^u_x \to F^s_y \times F^u_y.$$

Provided the splitting $F^s \oplus F^u$ is sufficiently close to $E^s \oplus E^u$, the map Γ_x is well defined by the formula

$$\Gamma_x(\mu_x) = (C_x + D_x\mu_x) \circ (A_x + B_x\mu_x)^{-1};$$

and given $\varepsilon > 0$ we may assume $\|C_x\|$ and $\|B_x\|$ so small that $L(\Gamma_x) \le L(D_x) L(A_x^{-1}) + \varepsilon \le ab + \varepsilon = \kappa$. Choose ε so that $\kappa < 1$ and $\kappa C^\alpha < 1$.

Let $L(F^s, F^u)$ be the vector bundle over M whose fiber over x is L_x; let D be its unit ball bundle. Then $\Gamma: D \to D$ is a bundle map covering f^{-1}; and Γ is C^1 if f is C^2. Moreover $L(\Gamma_x) \le \kappa < 1$. Clearly $L(f^{-1}) \le \max\{a, c\} = c$. By 6.1(b) the unique Γ-invariant section of D, which is λ, is α-Hölder; this proves 6.3(a). The proof of 6.3(b) follows similarly from 6.1(c). The proof of 6.3(bii) is left to the reader.

Let q be the fiber dimension of the bundle E^s, let $G_q(X)$ be the bundle of q-planes in T_XM for any subset $X \subset M$. Let $\theta: \Lambda \to G_q(\Lambda)$ assign E^s_x to $x \in \Lambda$.

6.4. THEOREM. *Let $\Lambda \subset U$ be a compact hyperbolic set for the C^1 embedding $M \supset U \xrightarrow{f} M$.*

(a) *If $abc^\alpha < 1, 0 < \alpha \le 1$, then $\theta: \Lambda \to G_q(\Lambda)$ is α-Hölder.*

(b) *Let $\{V_i\}$ be a collection of C^1 submanifolds of Λ such that $\bigcup_i V_i$ is invariant under f. If f is C^2 and $abc < 1$, then each map $\theta|V_i: V_i \to G_q$ is C^1.*

PROOF. Almost identical to that of 6.3 and left to the reader.

Now let $\{W^s_x\}_{x\in\Lambda}$ be the stable manifold system for Λ, and put $W^s = \bigcup_{x\in\Lambda} W^s_x$. Then $f(W^s) \subset W^s$ (assuming the metric on M adapted to Λ; or f could be replaced by f^n for n sufficiently large). Define $\text{glob}(W^s) = \bigcup_{n\ge 0} f^{-n}(W^s)$. Equivalently, $\text{glob}(W^s) = \bigcup_{x\in\Lambda}\{y \in U | \lim_{n\to\infty} d(f^ny, f^nx) \to 0\} = \bigcup_{x\in\Lambda}\text{glob}(W^s_x)$ where

$$\text{glob}(W^s_x) = \bigcup_{n\ge 0} f^n(W^s_x).$$

Each set $\text{glob}(W^s_x)$ is a disjoint union of q-dimensional submanifolds of M, and so is $\text{glob}(W^s)$. Define $\theta: \text{glob}(W^s) \to G_q$ by $\theta(y) = Tf^{-n}T_{f^ny}(W^sx)$ if $y \in f^{-n}W^s_x$, $x \in \Lambda$, $n \ge 0$. Then θ is well defined.

It may happen that W^s, and hence $\text{glob}(W^s)$, is open in M. This is the case for the 1-dimensional attractors of R. F. Williams [17].

6.5. THEOREM. *Assume W^s open in M. Let f be C^2. Then $\theta: \text{glob}(W^s) \to G_q(M)$ is C^1 provided $abc < 1$. In particular this is the case if Λ is a 1-dimensional attractor.*

PROOF. It suffices to prove that θ is C^1 in some neighborhood N of Λ in W^s; for if x is any point of $\text{glob}(W^s)$ there exists $n \ge 0$ such that $f^n(x) \in N$, and $\theta = \theta \circ f^n$.

Let $\tilde{E}^u \subset T_NM$ be a subbundle over a neighborhood $N \subset W^s$ of Λ extending E^u; we do not assume $\tilde{E}^u$ invariant. Let $\tilde{E}^s \subset T_NM$ be the subbundle whose fiber over y is $\theta(y)$. Choose N so that $f(N) \subset N$ (see 8.4 below). Choose $\varepsilon > 0$ so small that $(a + \varepsilon)(b + \varepsilon)(c + \varepsilon) < 1$. We may choose N so small that

$$\|Tf^{-1}|\hat{E}^u\| < a + \varepsilon, \qquad \|Tf|\hat{E}^s\| < b + \varepsilon, \qquad \|Tf|E^u\| < c + \varepsilon.$$

(For example given any compact N_0 such that $fN_0 \subset N_0$ and $\bigcap_{n \geq 0} f^n(N_0) = \Lambda$, let $N = f^p(N_0)$ for a large value of p.) Now the proof of 6.3(b) may be applied, replacing 6.1(c) by 6.2. We leave the details to the reader.

REMARK. In his paper [16], Williams assumed as an axiom that if Λ is a one dimensional attractor then the local projections $W^s \to \Lambda$, defined locally by mapping W^s_x to x, are Lipschitz. It is easy to see that this is in fact a consequence of Theorem 6.5 if f is C^2.

7. **Perturbations of hyperbolic sets.** We continue the standing hypothesis $M \supset U \xrightarrow{f} M$ is a C^k embedding and $\Lambda \subset U$ is a compact invariant hyperbolic set of skewness $\tau < 1$. For simplicity we assume the metric on M adapted to Λ.

7.1. THEOREM. *Let $\varepsilon > 0$. There exists a neighborhood $V \subset U$ of Λ and a neighborhood $\mathcal{N} \subset C^k(U, M)$ such that if $g \in \mathcal{N}$ then any invariant set of g in V is hyperbolic of skewness σ with $|\tau - \sigma| < \varepsilon$.*

If $K \subset V$ is compact, then $\bigcap_{n \in \mathbf{Z}} g^n K$ is the unique maximal invariant subset of K. Hence

7.2. COROLLARY. *In Theorem 7.1, every compact subset $K \subset V$ contains a unique maximal g-hyperbolic subset; this subset contains every g-invariant subset of K. In particular every compact g-invariant subset of V is hyperbolic.*

PROOF OF THEOREM 7.1. Let $W \subset U$ be a neighborhood of Λ over which the invariant splitting $E^u \oplus E^s$ of TM can be extended to a splitting $E_1 \oplus E_2$ of $T_W M$. Let $V \subset W$ be a neighborhood of Λ so small that if $x \in W$ then $T_x f^n$ is represented by a matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} : (E_1 \oplus E_2)_x \to (E_1 \oplus E_2)_{fx},$$

satisfying

$$\max\{\|A^{-1}\|, \|D\|\} < \tau + \varepsilon/2, \qquad \max\{\|B\|, \|C\|\} < \varepsilon/2,$$

where ε is as in Theorem 4.8. Let $\mathcal{N}$ be a neighborhood of f in $C^1(U, M)$ so small that if $g \in \mathcal{N}$ and $x \in V$ then $T_x g$ is represented by a matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

satisfying (a) and (b) of Proposition 4.10. By that Proposition any g-invariant in V is hyperbolic.

The next result shows that maximal hyperbolic sets enjoy a type of structural stability.

7.3. THEOREM. *Let $\Lambda \subset U$ be a compact hyperbolic set for the C^1 embedding $M \supset U \xrightarrow{f} M$. Given $\varepsilon > 0$ there is a compact neighborhood $B \subset U$ of Λ and a neighborhood $\mathcal{N}$ of f in $C^1(U, M)$ with the following properties: if $g_i \in \mathcal{N}$ for $i = 1, 2$, then g_i has a unique maximal hyperbolic set $\Lambda_i \subset V$ containing every invariant set of g_i in V; and there is a unique homeomorphism $h_1 : \Lambda_1 \to \Lambda_2$ such that $h_1 g_1 h_1^{-1} = g_2|\Lambda_2$; and $d(h_1, 1) \leq \varepsilon$. Moreover h_1 depends continuously on $(g_1, g_2) \in \mathcal{N} \times \mathcal{N}$.*

PROOF. Choose a compact neighborhood $V \subset U$ of Λ and a neighborhood $\mathcal{N}$ of f in $C^1(U, M)$ as in 7.1 and 7.2. Given g_1 and g_2 in $\mathcal{N}$, let Λ_1 and Λ_2 be their respective compact maximal hyperbolic sets in V. Let $\mathscr{C}(\Lambda_1, M)$ be the Banach manifold of continuous maps $h : \Lambda_1 \to M$. Define $\Phi : \mathscr{C}(\Lambda_1, U) \to \mathscr{C}(\Lambda_1, M)$ by $\Phi(h) = g_2 \circ h \circ g_1^{-1}$.

Let $i_1 : \Lambda_1 \to M$ be the inclusion of Λ_1. If V and $\mathcal{N}$ are sufficiently small, depending only on f, then Proposition 4.8 shows that the derivative of Φ at i_1 will be hyperbolic. Moreover Φ has a unique fixed point h_1 by Theorem 3.2. Therefore $g_2 h_1 = h_1 g_1$. Clearly $h_1(\Lambda_1)$ is g_2-invariant, and so $h_1(\Lambda_1) \subset \Lambda_2$. Similarly there exists $h_2 : \Lambda_2 \to \Lambda_1$ such that $g_1 h_2 = h_2 g_1$.

Observe that $h_1 h_2 g_2 = h_1 g_1 h_2 = g_2 h_1 h_2$. Therefore, by uniqueness, $h_1 h_2$ = identity map of Λ_2. Similarly $h_2 h_1$ = identity map of Λ_1; so h_1 and h_2 are homeomorphisms. The continuity of h_1 in (g_1, g_2) depends on the universal estimates for continuity of fixed points of contractions (see 1.1) and is left to the reader.

The next theorem means that the stable manifolds of a hyperbolic set move only slightly under perturbations.

7.4. THEOREM. *Referring to Theorem 7.3, let f be C^k. Let the stable manifold system of Λ be of size β. Then the neighborhoods $V \subset U$ of Λ and $\mathcal{N} \subset C^1(U, M)$ can be chosen so that if $g_1, g_2 \in \mathcal{N} \cap C^k(U, M)$, then the following conditions hold. Let $h : \Lambda_1 \to \Lambda_2$ be as in 7.3; let $x \in \Lambda_1$ and put $h(x) = y \in \Lambda_2$. Λ_i has a stable manifold system $\{W_x^i\}_{x \in \Lambda_i}$ for g_i of size $\geq \beta - \varepsilon$. Moreover if $x \in \Lambda_1$ and $h(x) = y \in \Lambda_2$, where $h : \Lambda_1 \to \Lambda_2$ is as in 7.3, there is a C^k diffeomorphism $\theta_x : W_x^1 \to W_y^2$ which is ε-close to the inclusion $W_x^1 \to M$ in $C^k(W_x^1, M)$. Moreover θ_x depends continuously on (g_1, g_2).*

PROOF. We assume familiarity with the proof of 3.2. It suffices to prove the theorem with $g_1 = f$. Put $g = g_2$. Let $\mathscr{C}(\Lambda, U)$ be the Banach manifold of continuous maps $h : \Lambda \to U$. Define $\Phi_g : \mathscr{C}(\Lambda, U) \to \mathscr{C}(\Lambda, M)$ by $\Phi_g(h) = ghf^{-1}$. Then Φ_g has the unique hyperbolic fixed point $h_g : \Lambda \to \Lambda_2$. If $\mathscr{W}_g \subset \mathscr{C}(U, M)$ is the stable manifold for Φ_g, it can be shown, imitating the proof of 3.2, that $W_{y,g} = \{h(x) | h \in \mathscr{W}_g\}$ is the stable manifold through $y = h(x) \in \Lambda_2$ for g. Now Φ_g depends continuously on g respecting the C^k topologies. The stable manifold function of Φ_g is a C^k map $\mathscr{G}_g : \mathscr{C}(E^s) \to \mathscr{C}(E^u)$, where $T_\Lambda \mathcal{N} = E^s \oplus E^u$ and $\mathscr{C}$ denotes the Banach space of bounded continuous sections. Since $\mathscr{G}_g$ depends C^k continuously on g, so does its graph, which is $\mathscr{W}_g$. A C^k diffeomorphism $\theta : \mathscr{W}_f \to \mathscr{W}_g$ is defined by $\theta_g(u, \mathscr{G}_f(u)) = (u, \mathscr{G}_g(u))$. (Here we identify $\mathscr{C}(\Lambda, U)$ with $\mathscr{C}(E^s) \times \mathscr{C}(E^u)$

by exponential coordinates.) Define $\theta_x : W_x \to W_{y,g}$ by $\theta_x(v(x)) = \theta(v)(x)$, for $x \in \Lambda$ and $v \in \mathscr{W}_f$. The details are left to the reader.

REMARK. $\{W_x^f\}_{x\in\Lambda}$ and $\{W_y^g\}_{g\in\Lambda_2}$ are "continuous families" of C^k submanifolds. The proper way to state Theorem 7.4 is to define the concept of a *continuous family* of C^k diffeomorphisms $\{\theta_x : W_x \to W_{h(x)}\}_{x\in\Lambda}$. We leave this task to the reader.

REFERENCES

1. Abraham and Robbin, *Transversal mappings and flows*, Benjamin, New York, 1967.
2. Arnold and Avez, *Problèmes ergodiques de la mecanique classique*, Gauthier-Villars, Paris, 1966.
3. Anosov, *Ergodic properties of geodesic flows . . .*, Soviet Math. 3 (1962), 1068–1070.
4. Dieudonné, *Foundations of modern analysis*, Academic Press, New York, 1960.
5. Artin and Tate, *Class field theory*, Benjamin, New York, 1967.
6. Hartman, *Ordinary differential equations*, Wiley, New York, 1964, Chapter IX.
7. Hirsch and Pugh, *Stable manifolds for hyperbolic sets*, Bull. Amer. Math. Soc. 75 (1969), 149–152.
8. Holmes, *A formula for the spectral radius of an operator*, Amer. Math. Monthly 75 (1968), 163–166.
9. Lang, *Introduction to differentiable manifolds*, Interscience, New York, 1962.
10. Mather, *Characterization of Anosov Diffeomorphisms*, Nederl. Akad. van Wetensch. Proc. Ser. A, Amsterdam 71 = Indag. Math. 30 (1968), no. 5.
11. Riesz and Nagy, *Leçons d'analyse functionelle*, Budapest, 1953.
12. Smale, *Differentiable dynamical systems*, Bull. Amer. Math. Soc. 73 (1967), 747–817.
13. ———, *The Ω-stability theorem*, these Proceedings, vol. 14.
14. ———, *Stable manifolds for differential equations and diffeomorphisms*, Ann. Scuol. Norm. Sup. Pisa 17 (1963), 97–116.
15. Sternberg, *Local contractions and theorem of Poincaré*, Amer. J. Math. 79 (1957), 809–824.
16. Williams, *One dimensional nonwandering sets*, Topology 6 (1967), 473–487.
17. ———, *Classification of one dimensional attractors*, these Proceedings, vol. 14.

UNIVERSITY OF CALIFORNIA, BERKELEY