An Analytic Proof of the Shadowing Lemma

By

K. R. Meyer* and George R. Sell*

(University of Cincinnati and University of Minnesota, U. S. A.)

1. Introduction

The shadowing lemma is one of the fundamental technical results used in the study of axiom A dynamical systems and homoclinic phenomena. The idea first appears in Anosov's [1] studies of geodesic flows and was later formalized by Bowen [2] for axiom A systems. These early results were established using detailed information about the intersection of the stable and unstable manifolds associated to hyperbolic sets. Robinson [7] gave a simple geometric proof of this lemma along with a variety of interesting applications. The reader is referred to Robinson's paper for a more complete introduction and further references.

In this paper we shall give a simple analytic proof of this lemma along with some applications. This simple proof facilitates a generalization of the shadowing lemma to skew product flows which is given in section IV. For discrete dynamical systems the shadowing lemma is as follows. Let $f: M \to M$ be a C^1 diffeomorphism of a smooth manifold M endowed with some Riemannian metric. An invariant set $\Lambda \subset M$ for f is called hyperbolic if there is a continuous splitting $TM|\Lambda = E^u \oplus E^s$ and constants $C > 0$, $0 < \mu < 1$ such that

$$(1) \qquad \|Df^n(x)(v)\| \leqslant C\mu^n \|v\| \quad \text{and} \quad \|Df^{-n}(x)(u)\| \leqslant C\mu^n \|u\|$$

for all $x \in \Lambda$, $v \in E^s_x$, $u \in E^u_x$ and $n \geqslant 0$. For $\alpha > 0$ an α-pseudo-orbit for $f|\Lambda$ is a bisequence $\{x_i\}_{i=-\infty}^{\infty}$ of points $x_i \in \Lambda$ such that $d(f(x_{i-1}), x_i) < \alpha$ for all i. One says that a f-orbit $\{f^i(y)\}$, β-shadows $\{x_i\}$ if $d(f^i(y), x_i) < \beta$ for all i.

The Shadowing Lemma. *If Λ is an compact, hyperbolic invariant set for $f: M \to M$, then for every $\beta > 0$ there is an $\alpha > 0$ such that every α-pseudo-orbit for $f|\Lambda$ is β-shadowed by some f-orbit $\{f^i(y)\}$. Moreover, there is a $\beta_0 > 0$ such that if $0 < \beta < \beta_0$ the f-orbit given above is uniquely determined by the α-pseudo-orbit.*

This lemma will be proved in the next section. After a brief discussion of some applications, it is generalized to almost periodic flows in section IV.

This paper is dedicated to the memory of Lamont Cranston and the lovely Margo Lane.

II. Discrete flows

In order to present the idea of the proof without notational complexities we shall assume that M is R^n. This allows us to work in a Banach space instead of a Banach manifold. Since $\Lambda \subset M$ is compact and we are interested in behavior near Λ, we may modify f outside a neighborhood of Λ so that all orbits of f are bounded. By averaging the metric we may assume that the constant C in (1) of the introduction is 1, see [6] for proof. Moreover, there is a neighborhood N of Λ such that the invariant splitting and the asymptotic estimates extend to this neighborhood, see [6]. Thus we shall assume that $f: R^n \to R^n$ is a C^1 diffeomorphism such that for each $x \in R^n$ the set $\{f^n(x)\}$ is bounded. Moreover, there is a neighborhood N of Λ with compact closure, a constant $0 < \mu < 1$ and a continuous splitting $TR^n | \overline{N} = E^s \oplus E^u$ such that if x and $f(x) \in \overline{N}$ then $\|Df(x)(u)\| < \mu \|u\|$ for all $u \in E^s_x$, and if x and $f^{-1}(x) \in \overline{N}$ then $\|Df^{-1}(x)(v)\| < \mu \|v\|$ for all $v \in E^u_x$.

Let $\mathcal{M} = \ell_\infty$ be the Banach space of bounded, bisequences $x = \{x_i\} = (\cdots, x_{-1}, x_0, x_1, \cdots)$, $x_i \in R^n$, with norm $\|x\| = \sup_i \|x_i\|$. Consider the map $\mathcal{F}: \mathcal{M} \to \mathcal{M}$ defined by $(\mathcal{F}(x))_i = f(x_{i-1})$. Note, that the bisequence $x = \{x_i\}$ is an f-orbit if and only if $\mathcal{F}(x) = x$, because $\mathcal{F}(x) = x$ means $f(x_{i-1}) = x_i$ for all i. Let $\mathcal{L} = \{x \in M : x_i \in \Lambda\}$ and $\mathcal{N} = \{x \in M : x_i \in N\}$. Thus $\mathcal{N}$ is a neighborhood of $\mathcal{L}$ in $\mathcal{M}$. Since Λ is compact there is a $\delta_0 > 0$ such that the δ_0-neighborhood of Λ lies in N. Even though it is not compact, the δ_0 neighborhood of $\mathcal{L}$ lies in $\mathcal{N}$.

Also note that a bisequence $x = \{x_i\}$ is an α-pseudo-orbit for f if and only if $\|\mathcal{F}(x) - x\| < \alpha$ and that $y = \{y_i\}$ β-shadows x if and only if $\mathcal{F}(y) = y$ and $\|y - x\| < \beta$. Thus, in this context the Shadowing Lemma says: *For every $\beta > 0$ there is a $\alpha > 0$ such that if $x \in \mathcal{L}$ and $\|\mathcal{F}(x) - x\| < \alpha$ then there is a $y \in \mathcal{N}$ such that $\mathcal{F}(y) = y$ and $\|x - y\| < \beta$.* In other words, if $\|\mathcal{F}(x) - x\| < \alpha$ then $\mathcal{F}$ has a fixed point y in an α-neighborhood of x. This formulation suggest the use of the inverse function theorem with the standard estimate on the domain of the inverse function.

The function $\mathcal{F}$ is differentiable with derivative given by

$$(D\mathcal{F}(x)(w))_i = Df(x_{i-1})(w_{i-1}),$$

where $x, w \in \mathcal{M}$. Let $x \in \mathcal{N}$ and define $\mathcal{S}_x = \times_{-\infty}^{\infty} E^s_{x_i} \cap \mathcal{M}$, $\mathcal{U}_x = \times_{-\infty}^{\infty} E^u_{x_i} \cap \mathcal{M}$. Since the original splitting is continuous, both $\mathcal{S}_x$ and $\mathcal{U}_x$ are closed, linear complimentary subspaces of $\mathcal{M}$ for each $x \in \mathcal{N}$, thus $\mathcal{M} = \mathcal{S}_x \oplus \mathcal{U}_x$. Also, $D\mathcal{F}(x)$ maps $\mathcal{S}_x$ into $\mathcal{S}_{\mathcal{F}(x)}$ and $\mathcal{U}_x$ into $\mathcal{U}_{\mathcal{F}(x)}$.

Using these facts we shall show that there is a constant K such that

$$(2) \qquad \|D\mathcal{F}(x) - I\| \leqslant K \quad \text{and} \quad \|(D\mathcal{F}(x) - I)^{-1}\| \leqslant K$$

for all $x \in \mathcal{N}$. The first estimate follows the fact that Df is uniformly bounded

on the compact set $\bar{N}$. Because of linearity, the second estimate needs only be proved on the two complimentary subspaces $\mathcal{S}_x$ and $\mathcal{U}_x$. For $u \in \mathcal{S}_x$ we have

$$\|D\mathcal{F}(x)(u) - u\| \geqslant \|u\| - \|D\mathcal{F}(x)(u)\|$$

$$\geqslant \|u\| - \mu\|u\| = (1 - \mu)\|u\|$$

Hence $\|D\mathcal{F}(x) - I)|_{\mathcal{S}_x}^{-1}\| \leqslant (1 - \mu)^{-1}$.

Since Df is invertible on the compact set $\bar{N}$ there is a constant $K_3 > 0$ such that $\|Df(x)(w)\| \geqslant K_3\|w\|$ for all $x \in \bar{N}$ and $w \in \mathbf{R}^n$. In particular this implies that $\|D\mathcal{F}(x)(v)\| \geqslant K_3\|v\|$ for all $x \in N$ and $v \in \mathcal{U}_x$. Let $v \in \mathcal{U}_x$, then

$$\|D\mathcal{F}(x)(v) - v\| = \|D\mathcal{F}(x)\{v - \mathcal{F}^{-1}(x)(v)\}\|$$

$$\geqslant K_3(\|v\| - \mu\|v\|) = K_3(1 - \mu)\|v\|$$

Hence $\|(D\mathcal{F}(x) - I)|_{\mathcal{U}_x}^{-1}\| \leqslant K_3^{-1}(1 - \mu)^{-1}$, which implies the second estimate in (2).

The following version of the inverse function theorem with estimate can be found in [4]. Let $\mathcal{B}$ be a Banach space, $B_\delta(x)$ the ball of radius δ about $x \in \mathcal{B}$, $\mathcal{G}: B_\delta(x) \to \mathcal{B}$ a C^1 function and $y_0 = \mathcal{G}(x_0)$.

Theorem. *Assume $D\mathcal{G}(x)$ has a bounded inverse for all $x \in B_\delta(x_0)$ and*

$$\|D\mathcal{G}(x)\| \leqslant K \quad and \quad \|D\mathcal{G}^{-1}(x)\| \leqslant K$$

for all $x \in B_\delta(x_0)$ where K is a constant. Let $p = \delta/K^2$ and $q = \delta/K$, then there exists a domain Ω, $B_p(x_0) \subset \Omega \subset B_\delta(x_0)$ such that $\mathcal{G}$ is one-to-one on Ω. Moreover, $B_q(y_0) \subset \mathcal{G}(B_\delta(y_0))$.

Apply this form of the inverse function theorem to the function $\mathcal{G} = \mathcal{F} - I$ where I is the idenity map of $\mathcal{B}$. As remarked before there is a $\delta_0 > 0$ such that $B_{\delta_0}(x) \subset N$ for all $x \in \mathcal{L}$. Let $\alpha < \delta_0$ and so the estimates (1) on the derivative of $\mathcal{F}$ hold on $B_\alpha(x)$ for $x \in \mathcal{L}$. Define $\beta = \alpha/K$. Then the inverse function theorem gives $\mathcal{G}(B_\alpha(x)) \subset B_\beta(\mathcal{G}(x))$. So, if $\|\mathcal{G}(x)\| = \|\mathcal{F}(x) - x\| < \beta$ then $0 \in B_\beta(\mathcal{G}(x))$ or $0 \in \mathcal{G}(B_\alpha(x))$. Thus there is a $y \in B_\alpha(x)$ such that $\mathcal{G}(y) = \mathcal{F}(y) - y = 0$. Thus y is an f orbit that α-shadows the β-pseudo-orbit x.

If we take $\beta_0 = \delta_0/K^2$ then the inverse function theorem yields that $\mathcal{G}$ is one-to-one on $B_{\beta_0}(x)$ for $x \in \mathcal{L}$. Thus the distance between zeros of $\mathcal{G}$ is at least β_0 and this proves the uniqueness part of the Shadowing Lemma.

In the general case when M is a differentiable manifold of dimension n, the space $\mathcal{M}$ is a Banach manifold modeled on the Banach space of bounded bisequences in $\mathbf{R}^n$ with sup norm. The estimates given above hold uniformly in each coordinate patch. Consequently the proof carries over with little or no change.

III.   Remarks

From the uniqueness part of the above proof it follows that $\mathscr{G}$ is one-to-one on B_{β_0} for all $x \in \mathscr{L}$. That means that if x, $y \in \mathscr{L}$ and $x \neq y$, $\mathscr{F}(x) = x$, $\mathscr{F}(y) = y$ then $\|x - y\| > \beta_0$ or if x_0, $y_0 \in \Lambda$ and $x_0 \neq y_0$ then for some n, $\|f^n(x_0) - f^n(y_0)\| > \beta_0$. Thus $f|\Lambda$ is expansive with expansive constant β_0.

The original proof [1] of the structural stability of Anosov diffeomorphisms used the local geometry of the stable and unstable manifolds to establish the shadowing lemma and the shadowing lemma was used to construct the homeomorphism. Subsequently more analytic proofs were found [5].

The simplified proof above gives an alternate short proof of the structural stability theorem for Anosov diffeomorphisms. Recall that an Anosov diffeomorphism is a smooth diffeomorphism f of a compact manifold M which has a hyperbolic structure on all of M. An outline of the proof of the structural stability of Anosov diffeomorphism based on the shadowing lemma is as follows: The proof of the shadowing lemma given above depends entirely on the estimates on the derivatives of the associated map $\mathscr{F}: \mathscr{M} \to \mathscr{M}$ given above. By continuity considerations if g is a diffeomorphism that is C^1 close to f then similar estimates hold for $\mathscr{G}$ (the lift of g to $\mathscr{M}$) and so g has the shadowing property also. If g is within α of f in the C^1 topology then for each $x \in M$, $\{f^n(x)\}$ is a α-pseudo-orbit for g and so by the shadowing lemma there is a unique $y = h(x) \in M$ such that $d(f^n(x), g^n(y)) < \beta$ for all n. Since the shadowing lemma was established using the implicit function theorem, $y = h(x)$ depends continuously on x for all $x \in M$. To assure that h is one-to-one let $\beta < \beta_0/3$. Let $x_1 \neq x_2$, $y_1 = h(x_1)$, $y_2 = h(x_2)$. By the expansive property discussed above there is an N such that $d(f^N(x_1)$, $f^N(x_2)) > \beta_0$, but $d(f^N(x_i), g^N(y_i)) < \beta_0/3$ so $d(g^N(y_1), g^N(y_2)) > \beta_0/3$ or $y_1 \neq y_2$.

Thus $h: M \to M$ is a continuous, one-to-one mapping of a compact Hausdorff space and so is a homeomorphism into M. $k(M)$ is closed in M and from general topology h is a local homeomorphism so $h(M)$ is open in M. Thus h is onto and hence a homeomorphism. Since $y = h(x)$ and $d(f^n(x), g^n(y)) < \alpha$ for all n, we have

$$d(f^{n-1}(f(x)), g^{n-1}(g(y))) = d(f^n(x), g^n(x)) < \alpha \leqslant \beta_0$$

Thus, the f orbit through $f(x)$ is β_0-shadowed by the g-orbit through $g(y)$, so by the uniqueness $h(f(x)) = g(y) = g(h(x))$. This means $h \circ f = g \circ h$ or h carries f-orbits into g-orbits.

IV.   Non-autonomous differential equations

Let $\mathscr{C} = \mathscr{C}(\mathbf{R}^1 \times \mathbf{R}^n, \mathbf{R}^n)$ be the space of all continuous functions from $\mathbf{R}^1 \times$

R^n into R^n endowed with the topology of uniform convergence on compact sets (the compact open topology). Let τ be a real number, then the τ-translate of $J \in \mathscr{C}$ is the function $J_\tau \in \mathscr{C}$ defined by $J_\tau(t, x) = J(t + \tau, x)$. The function $\pi: R \times \mathscr{C} \to \mathscr{C}: (\tau, J) \to J_\tau$ defines a flow on $\mathscr{C}$, see [9]. Let $N \subset \mathscr{C}$ be a compact invariant set in $\mathscr{C}$. By Ascoli's theorem, a subset of $\mathscr{C}$ is compact if it is closed and the members are uniformly bounded and equi-continuous on compact subsets of $R^1 \times R^n$. An important special case is the hull of $J \in \mathscr{C}$ where the hull of J, denoted by $H = H(J)$, is the orbit closure of J under the flow defined by π. The hull of J is compact if and only if J is bounded and uniformly continuous on sets of the form $R^1 \times K$ where K is compact in R^n.

Our first basic assumptions about N are:

i)   for each $f \in N$, $D_2 f$ is continuous and uniformly bounded on sets of the form $R^1 \times K$ where K is compact in R^n

ii) for each $f \in N$, all solutions of the equation

(1) $$\dot{x} = f(t, x)$$

are bounded for all t. Again this is just a technical assumption.

As a simple example consider the linear equation $\ddot{\xi} + 4\xi = \sin t + \sin \sqrt{2}\, t$ written as a system $\dot{\xi}_1 = \xi_2$, $\dot{\xi}_2 = -4\xi_1 + \sin t + \sin \sqrt{2}\, t$. Here we take $F(t, \xi_1, \xi_2) = (\xi_2, -4\xi_1 + \sin t + \sin \sqrt{2}\, t)$ and N the hull of F. It is easy to see that the hull of F is $\{G = (\xi_2, -4\xi_1 + \sin (t + \alpha) + \sin (\sqrt{2}\, t + \beta)$ where α and β are real constants defined mod $2\pi\}$ and thus the hull is homeomorphic to T^2.

For $f \in N$, let $\phi(t, x_0, f)$ be the solution of (1) which satisfies $\phi(0, x_0, f) = x_0$ then $\pi: R^1 \times R^n \times N \to R^n \times N: (t, x_0, f) \to (\phi(t, x_0, f), f_t)$ defines a flow, see [9]. Let $M \subset R^n$ be compact and such that $M \times N$ is invariant under π. We say that $M \times N$ has a hyperbolic structure (normal to N) if for each $(x_0, f) \in M \times N$ the variational equation of (1) along the solution $\phi(t, x_0, f)$ has an exponential dichotomy which is uniform on $M \times N$. The precise meaning of the above is as follows.

The variational equation of (1) along $\phi(t, x_0, f)$ is

(2) $$\dot{y} = A(t)y$$

where $A(t) = D_2 f(t, \phi(t, x_0, f))$. Let $Y(t) = Y(t, x_0, f)$ be the fundamental matrix solution of (2) that satisfies $Y(0) = I$. Then (2) has an exponential dichotomy at (x_0, f) if there exists a projection matrix P and positive constants K and α such that

(3) $|Y(t)PY^{-1}(s)| \leqslant Ke^{-\alpha(t-s)} \qquad$ for $\quad t \geqslant s$

$|Y(t)(I - P)Y^{-1}(s)|\,| \leqslant Ke^{-\alpha(t-s)} \qquad$ for $\quad s \geqslant t$.

The matrix A and hence Y, P, K and α depend on $(x_0, f) \in M \times N$. It is known [8]

that P depends continuously on (x_0, f). A simple uniformity argument shows that K and α can be chosen independently of (x_0, f) provided (x_0, f) remains in a compact set.

As an example consider $\dot{\xi}_1 = \xi_2$, $\dot{\xi}_2 = -\xi_2 + (-4 + \varepsilon \sin(t + \alpha) + \varepsilon \sin(\sqrt{2}t + \beta))\xi_1$ where α and β are constants defined mod 2π and ε is a small parameter. The origin is a solution and for ε sufficiently small the origin is uniformly exponentially stable. In this case $P = I$ and $M = \{0\}$. Of course all solutions of the above equations may not be bounded for all t. Since we are interested in a neighborhood of $M \times N$ we can modify the equation in this example outside a large sphere so that the solutions are bounded for all time.

We will next show how the shadowing lemma applies to (1). Let $\mathscr{CB}^k = \mathscr{CB}^k(R^1, R^n)$ denote the space of functions $f: R^1 \to R^n$ with k continuous and bounded derivatives with norm defined by $\|f\|_k = \sup_{0 \leqslant s \leqslant k} \sup_{t \in R} \|D^s f(t)\|$. Define $\mathscr{F}: \mathscr{CB}^1 \times N \to \mathscr{CB}^0: (\phi, f) \to \dot{\phi} - f \circ \phi$. The function $\phi \in \mathscr{CB}^1$ is a solution of (1) if and only if $\mathscr{F}(\phi, f) = 0$. In this case $(\phi(t), f_t)$ is an orbit of π. The function $\mathscr{F}$ is differential with respect to the first argument and $D_1 \mathscr{F}(\phi, f)(\psi) = \dot{\psi} - (Df \circ \phi)\psi$.

For $\alpha > 0$ an α-pseudo-orbit is a pair $(\psi, g) \in \mathscr{CB}^1 \times N$ such that $\|\mathscr{F}(\psi, g)\|_0 < \alpha$. For $\beta > 0$ we say that an orbit (ϕ, f), β-shadows the pseudo-orbit (ψ, f) if $\|\phi - \psi\|_1 < \beta$ (Note that both the orbit and the pseudo-orbit have the same second argument. This means ϕ and ψ are solutions of the same differential equation).

The Shadowing Lemma. *Let M be a compact, hyperbolic invariant set for π. Then for every $\beta > 0$ there is an $\alpha > 0$ such that every α-pseudo-orbit in M is β-shadowed by an orbit in M. Moreover, there is a $\beta_0 > 0$ such that if $0 < \beta < \beta_0$ the orbit given above is uniquely determined by the pseudo-orbit.*

The proof of this variation of the shadowing lemma follows the same outline as the proof of the shadowing lemma for diffeomorphisms with some minor modifications. First the $\mathscr{F}$ for diffeomorphisms in section 2 had fixed points for f orbits whereas the $\mathscr{F}$ in this section has zeros for orbits. Thus here we will have to estimate $D_1 \mathscr{F}$ and its inverse. The estimation of $D_1 \mathscr{F}$ is straightforward and the estimation of $D_1 \mathscr{F}^{-1}$ follows from

Theorem. *Let $A(t)$ be a continuous and bounded $n \times n$ matrix for all t and $p \in \mathscr{CB}^0$, then the equation*

$$(4) \qquad\qquad \dot{y} = A(t)y + p(t)$$

has a unique solution in $\mathscr{CB}^1$ if and only if the homogeneous equation $\dot{y} = A(t)y$ has an exponential dichotomy. Moreover, if we denote this solution by Lp then L is a bounded linear operator from $\mathscr{CC}^0$ to $\mathscr{CB}^1$ with norm which can be estimated in terms of α and K only (See Coppel [3]. Prop. 3.2.).

Solving the differential equation (4) with $p \in \mathscr{CB}^0$ for a solution in $\mathscr{CB}^1$ is the same as solving $D_1 \mathscr{F}(y) = p$. Thus, this classical theorem states that $L = (D_1 \mathscr{F})^{-1}$ exists as a bounded linear operator for all $(\phi, f) \in M \times N$ and has a uniform estimate.

This is not quite enough to use the inverse function theorem. We need that L is continuous, but this follows from the roughness of exponential dichotomies, see Coppel [3] Prop. 4.1. This theorem also shows that the estimates can be extended to a full neighborhood of $M \times N$ in $\mathbf{R}^n \times N$.

References

[1] Anosov, D., Geodesic flows on closed Riemann manifolds with negative curvature, Proc. Steklov Inst. Math., **90** (1967).

[2] Bowen, R., ω-limit sets for Axiom A diffeomorphisms, J. Differential Equations, **18** (1975), 333–339.

[3] Coppel, W. A., *Dichotomies in Stability Theory*, Lecture Notes in Math. 629, Springer-Verlag, 1978.

[4] Hartman, P., *Ordinary Differential Equations*, J. Wiley, New York, 1964.

[5] Moser, J., On a theorem of D. Anosov, J. Differential Equations, **5** (1969), 411–440.

[6] Nitecki, Z., *Differentiable Dynamics*, MIT Press, Cambridge Mass., 1971.

[7] Robison, C., Stability theorems and hyperbolicity in dynamical systems, Rocky Mountain J. Math., (1977), 425–437.

[8] Sacker, R. and Sell, G., Existence of dichotomies and invariant splitting for linear differential systems I, J. Differential Equations, **15** (1974), 429–458.

[9] Sell, G., *Topological Dynamics and Ordinary Differential Equations*, van Nostrand, London, 1971.

nuna adreso
K. R. Meyer
Department of Mathematical
Sciences
University of Cincinnati
Cincinnati, Ohio 45221
U. S. A.

G. R. Sell
Institute for Mathematics
and Its Applications
University of Minnesota
Minneapolis, Minnesota 55455
U. S. A.