RC 20506 (90887) 7/22/96
Computer Science/Mathematics 11 pages

Research Report

Stable Ergodicity and Stable Accessibility

Charles Pugh and Michael Shub
IBM Research Division
T.J. Watson Research Laboratory
P.O. Box 218
Yorktown Heights, NY 10598

LIMITED DISTRIBUTION NOTICE

This report has been submitted for publication outside of IBM and will probably be copyrighted if accepted for publication. It has been issued as a Research Report for early dissemination of its contents. In view of the transfer of copyright to the outside publisher, its distribution outside of IBM prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article (e.g., payment of royalties).

IBM Research Division
Almaden • T.J. Watson • Tokyo • Zurich

STABLE ERGODICITY AND STABLE ACCESSIBILITY

by

Charles Pugh and Michael Shub*

Abstract. In this paper we discuss how a little hyperbolicity goes a long way towards guaranteeing stable ergodicity. Examples to which our theory applies include translations on certain homogeneous spaces and the time one map of the geodesic flow for a manifold of constant negative curvature. We also discuss the role played by a concept from control theory, accessibility by paths tangent to a pair of subbundles of the tangent bundle.

1. **Introduction.** Let $f : M \to M$ be a C^2 diffeomorphism and assume that f preserves Lebesgue volume μ,

$$\mu(fA) = \mu A$$

for all measurable sets $A \subset M$. The diffeomorphism is **ergodic** if the only measurable sets it leaves invariant are trivial:

$$fA = A \quad \text{implies that either } \mu A = 0 \text{ or } \mu(A^c) = 0.$$

f is **stably ergodic** if it is ergodic and so is every volume preserving diffeomorphism that C^2-approximates it. Similar definitions make sense for flows. Anosov (1967) was the first to show that stably ergodic diffeomorphisms and flows exist. Specifically, his results imply that every area preserving diffeomorphism of the 2-torus that C^2-approximates the "cat map"

$$f = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \mod \mathbb{Z}^2$$

is ergodic. Hyperbolicity of f is a key ingredient of the proof. (By the way, it remains unknown to this day whether C^2 can be relaxed to C^1. See Bowen (~~19xx~~ 1995) and Robinson & Young (~~19xx~~ 1980).)

Boltzman's ergodic hypothesis — that ergodicity is a generic property of volume preserving dynamical systems — underlies statistical mechanics and much of physical thinking. It is encapsulated in the following picture of two gasses mixing.

---

* Partially supported by an NSF grant.

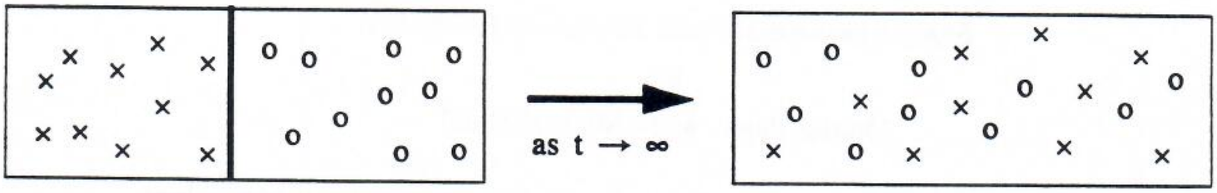

Figure 1. Gas ergodicity.

Yet, in 1954, Kolmogorov showed that Boltzman was wrong. The twist map

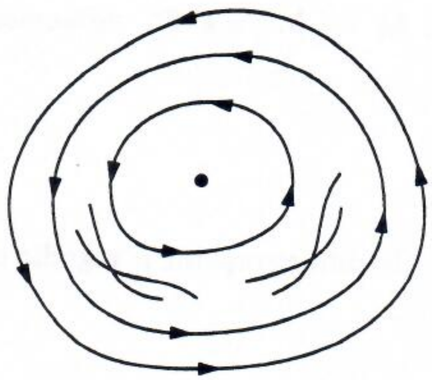

Figure 2. The twist map.

is an area preserving diffeomorphism of the plane that is not ergodic, nor is any area preserving diffeomorphism that C^4–approximates it, due to the persistence of the invariant KAM circles. Nevertheless, *it is our contention that Boltzman was more right than wrong:* in the joint presence of hyperbolicity and KAM dynamics, hyperbolicity can overwhelm the KAM dynamics and produce robust statistics in the form of ergodicity.

Consider the following four conditions that the diffeomorphism f may satisfy.

(1) f is partially hyperbolic.

(2) f is dynamically coherent.

(3) The stable and unstable bundles of f have the accessibility property.

(4) f has good holonomy.

See § 2 for details about (1) – (4). In our paper, Pugh & Shub (1996), we prove

Theorem A. Ergodicity of f is implied by (1), (2), (3), (4).

Theorem B. Stable ergodicity of f is implied by (1), (2), (3) and

(5) The partially hyperbolic splitting of f is C^1.

(6) Tf has sufficiently bunched spectrum.

Theorem C. The time one map of the geodesic flow on $T_1V = M$ is stably ergodic if V is a compact Riemann manifold with constant negative curvature.

Theorem D. Certain translations of homogeneous spaces are stably ergodic.

We view Theorem A as saying that although f may have completely anti-ergodic and KAM behavior on the center manifolds, the center dynamics can be *irrelevant* to ergodicity. Hyperbolicity in the normal directions can rule the scene. The proofs of Theorems C and D amount to verifying the hypotheses of Theorem B, while the proof of Theorem B amounts to verifying (1) - (4) for perturbations of f.

When V has negative curvature, Anosov showed that the geodesic flow φ on the unit tangent bundle $M = T_1V$ is ergodic as a flow. Any measurable set $A \subset M$ that is invariant under the flow is a zero set or the complement of a zero set. In fact, Anosov showed that this continues to be true for all flows that preserve volume and C^2-approximate φ; i.e., he established stable ergodicity of φ as a flow. Theorem C is stronger in two ways:

(a) Ergodicity of the time one map of a flow implies ergodicity of the flow but not conversely. Constant time suspensions give examples of this.

(b) Perturbations of the time one map of a flow need not even embed in a perturbation flow.

Together with Matt Grayson (1994) we proved Theorem C when V has dimension two and has constant negative curvature. In her thesis, Amie Wilkinson (1995) proved Theorem C when V has dimension two and has variable negative curvature. With certain technical improvements, our methods lead to a generalization of Theorem C to higher dimensional manifolds with variable negative pinched curvature. We gratefully acknowledge useful conversations with Jonathan Brezin about Theorem D, and with Eduardo Sontag about the control theory literature.

2. Details. We assume that M is compact, boundaryless, and m-dimensional. The volume μ, expressed in local coordinates, corresponds to an m-form $\delta(x)\, dx_1 \ldots dx_n$, where δ is smooth and positive.

The diffeomorphism $f : M \to M$ is **partially hyperbolic** if Tf leaves invariant a continuous splitting $TM = E^u \oplus E^c \oplus E^s$, where $E^u \neq 0 \neq E^s$, and, with respect to some fixed Riemann structure on TM, Tf expands E^u, Tf contracts E^s, and for all $p \in M$

$$\sup \|T_p^s f\| < \inf m(T_p^c f) \quad \text{and} \quad \sup \|T_p^c f\| < \inf m(T_p^u f).$$

$T^u f$, $T^c f$, $T^s f$ are the restrictions of Tf to E^u, E^c, E^s. The bundles, tangent maps, etc. are referred to as unstable, center, and stable respectively. The notation $m(T)$ refers to the **conorm** (or **minimum norm**) of a linear transformation T, $m(T) = \inf\{|Tv| : |v| = 1\}$.

According to Hirsch, Pugh, & Shub (1977), if f' C^1-approximates f then f' is also partially hyperbolic. It is also shown there that unique f-invariant foliations, $\mathcal{W}^u$ and $\mathcal{W}^s$, tangent to E^u and E^s, exist and their leaves are dynamically characterized according to strong backward and forward asymptoticity respectively.

In general there is no reason to expect that a center bundle integrates to an invariant foliation. This leads us to say that f is **dynamically coherent** if E^{cu}, E^c, E^{cs} integrate to f-invariant foliations $\mathcal{W}^{cu}$, $\mathcal{W}^c$, $\mathcal{W}^{cs}$, and

$$\mathcal{W}^u \text{ and } \mathcal{W}^c \text{ subfoliate } \mathcal{W}^{cu}, \text{ while } \mathcal{W}^c \text{ and } \mathcal{W}^s \text{ subfoliate } \mathcal{W}^{cs}.$$

(One foliation **subfoliates** a second if each leaf of the second is a union of leaves of the first.) The phrase "dynamically coherent" indicates that the unstable, center unstable, center, center stable, and stable orbit-classes fit together nicely.

Together with Moe Hirsch, ~~(in HPS)~~ we investigated normally hyperbolic invariant foliations and laminations. It is just a matter of unraveling the definitions to show that if a partially hyperbolic diffeomorphism leaves invariant a foliation $\mathcal{W}^c$ tangent to E^c then it is 1-normally-hyperbolic at $\mathcal{W}^c$, and conversely, if a diffeomorphism is 1-normally hyperbolic at an invariant foliation then it is partially hyperbolic.

Accessibility is a concept from control theory and is discussed in §3.

It remains to explain what "good holonomy" and "bunched spectrum" mean. At points p, q of the same unstable leaf W^u draw transversal discs and define a map from one transversal to the other by sliding along the leaves of W^u. This is an unstable holonomy map, say h. It is known that h is a biHölder homeomorphism, and that it is absolutely continuous with positive, continuous Radon Nikodym derivative. See Pugh & Shub (1972). Its Hölder exponent can be estimated. See Pugh, Shub, & Wilkinson (1996). Similarly the bundles E^u, E^c, E^s are Hölder and their Hölder exponents can be estimated. **Good holonomy** means that all Hölder exponents are greater than

$$\theta_m = \frac{\sqrt{100m^2 + 1} - 1}{10m}$$

where $m = \dim M$. This dimensional constant is greater than $9/10$ and increases to 1 as $m \to \infty$.

By the **spectrum** of Tf we mean the spectrum of the operator $\sigma \mapsto Tf \circ \sigma \circ f^{-1}$ defined on the space of bounded sections of TM. Spectral bunching conditions are used to prove that f has good holonomy. They say that the spectra of $T^u f$, $T^c f$, and $T^s f$ lie in thin, well separated annuli. More precisely, we assume that $a < b < c < 1 < d < e < g$ and

The spectrum of $T^s f$ lies in the annulus with radii a, b.
The spectrum of $T^c f$ lies in the annulus with radii c, d.
The spectrum of $T^u f$ lies in the annulus with radii e, g.

f has **sufficiently bunched spectrum** if $0 < \theta_m < \theta < 1$ and

$$d < \min\{ce, c/b, ea^\theta, eg^{-\theta}\} \quad \text{and} \quad b < \min\{ca^\theta, cg^{-\theta}\}.$$

3. Stable Accessibility. Accessibility is a concept from control theory. We follow the development presented by Grasse (1984), Lobry (1973), and Sussman (1976). See also Gromov (1995). Let E, F be subbundles of the tangent bundle TM of M. They need not be independent, nor need they span TM. An **(E, F) - path** is a piecewise C^1 path $\gamma : [a, b] \to M$ whose tangent lies alternately in E and F. That is, $[a, b]$ has a partition $a = t_0 < t_1 < \ldots < t_n = b$, and on the subintervals $I_i = [t_{i-1}, t_i]$ we have $(\gamma|_{I_1})' \in E$, $(\gamma|_{I_2})' \in F$, etc.

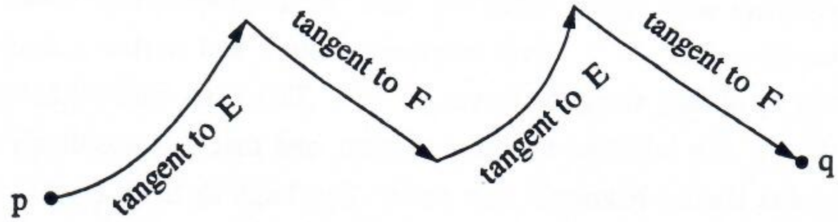

Figure 3. An (E, F) – path.

The pair (E, F) has the **accessibility property** if all p, q $\in$ M can be joined by (E, F) – paths with a fixed number L of legs, each leg having arclength < 1. The pair (E, F) has the **stable accessibility property** if for each nearby pair of subbundles (E', F') and all p, q $\in$ M, there is an (E', F') – path from p to q having $\leq$ L legs, each of arclength < 1.

A subbundle $E \subset TM$ is **uniquely integrable** (in the sense of control theory) if it can be spanned by continuous vector fields having unique integral curves. Clearly, if E is C^1 it is uniquely integrable. Unique integrability of E does not imply integrability in the Frobenius sense – E need not integrate to a foliation. A classic result in control theory is the

Stable Accessibility Theorem. Let E, F be C^1 subbundles of TM and assume (E, F) has the accessibility property. Then (E, F) has the stable accessibility property with respect to uniquely integrable approximations.

In order to comment on our use of this theorem and propose some extensions of it, we sketch the proof.

First, one chooses C^1 tangent vector fields $X_1, \ldots, X_k$ that span E and C^1 tangent vector fields $X_{k+1}, \ldots, X_{2k}$ that span F. (They need not be linearly independent.) Then one forms the C^1 **multiflow**

$$\Phi(\tau, x) = \varphi_{1, t_1} \circ \varphi_{k+1, t_{k+1}} \circ \ldots \circ \varphi_{k, t_k} \circ \varphi_{2k, t_{2k}}(x)$$

where $\tau = (t_1, \ldots, t_{2k})$, $x \in M$, and φ_i is the X_i–flow,

$$\left.\frac{d\varphi_{i,t}(x)}{dt}\right|_{t=0} = X_i(x).$$

By assumption, for every $p, q \in M$, there is a $\tau^* \in \mathbb{R}^{2k}$ such that $\Phi(\tau^*, p) = q$. In fact it turns out (read Grasse's paper to see why) that such a τ^* exists for which the C^1 map

$$\begin{aligned} \phi : \mathbb{R}^{2k} &\rightarrow M \\ \tau &\mapsto \Phi(\tau, p) \end{aligned}$$

has rank m at $\tau = \tau^*$. (The manifold M has dimension m.) This implies that there exists a smooth, compact m-dimensional disc D in $\mathbb{R}^{2k}$, centered at τ^*, such that ϕ carries D diffeomorphically onto a compact neighborhood U of q in M.

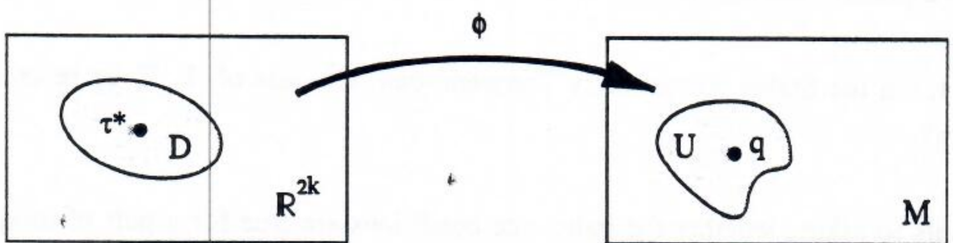

Figure 4. ϕ submerses over a neighborhood of q.

Next, one assumes that E', F' are uniquely integrable and approximate E, F in the uniform, C^0 sense. Then one approximates $X_1, \ldots, X_{2k}$ by uniquely integrable fields $X_1', \ldots, X_{2k}'$ that span E', F'. They generate a multiflow

$$\Phi'(\tau, x) = \varphi_{1}{}'_{,t_1} \circ \varphi_{k+1}{}'_{,t_{k+1}} \circ \cdots \circ \varphi_{k}{}'_{,t_k} \circ \varphi_{2k}{}'_{,t_{2k}}(x)$$

where φ_i' is the X_i'-flow. Note that Φ' is continuous, but it may fail to be C^1 because the fields X_i' need not be C^1. On compact subsets of $\mathbb{R}^{2k} \times M$, Φ' uniformly approximates Φ. Hence, the restriction of $\Phi'(\tau, p)$ to the disc D, call it ϕ', uniformly approximates ϕ.

Any point u interior to U has topological index one with respect to the restriction of ϕ to the boundary of D,

$$\operatorname{Index}(u, \phi|_{\partial D}) = 1 .$$

Topological index is homotopy invariant. Hence, there exists a smaller neighborhood V of q such that for all $u \in V$ and all ϕ' that uniformly approximate ϕ, $\text{Index}(u, \phi'|_{\partial D}) = 1$. This implies that $V \subset \phi'(D)$. Compactness of M completes the proof.

In our proof of Theorem B we are justified in using the Stable Accessibility Theorem because the hyperbolic bundles of the perturbation f' of f actually *are* uniquely integrable, although they are usually neither C^1 nor Lipschitz. The reasons for this are dynamical.

To extend Theorem B, by relaxing the hypothesis that E^u, E^s are C^1, it would be convenient to have a positive answer to

Question 1. In the Stable Accessibility Theorem, can C^1-ness of E, F be relaxed to unique integrability?

This amounts to asking whether the index one conditions are true for a pair of uniquely integrable bundles with the accessibility property. Allied to Question 1 is another question, one that we have no pressing need to answer, but a question that we find intriguing in its own right.

Question 2. Assume that E, F are C^∞ subbundles of TM, (E, F) has the accessibility property, and E', F' are continuous subbundles of TM that uniformly approximate E, F, but that are *not* necessarily uniquely integrable. Does (E', F') still have the accessibility property?

As a test case, consider $\mathbf{R}^3$ and three continuous vector fields X_1, X_2, X_3 that uniformly approximate $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$, but are not necessarily uniquely integrable. Let $p = (0, 0, 0)$, $q = (1, 1, 1)$, and ask: Do there exist an X_1-solution γ_1, an X_2-solution γ_2, and an X_3-solution γ_3 such that for some (t_1, t_2, t_3) near $(1, 1, 1)$, $\gamma_1(0) = p$, $\gamma_2(0) = \gamma_1(t_1)$,

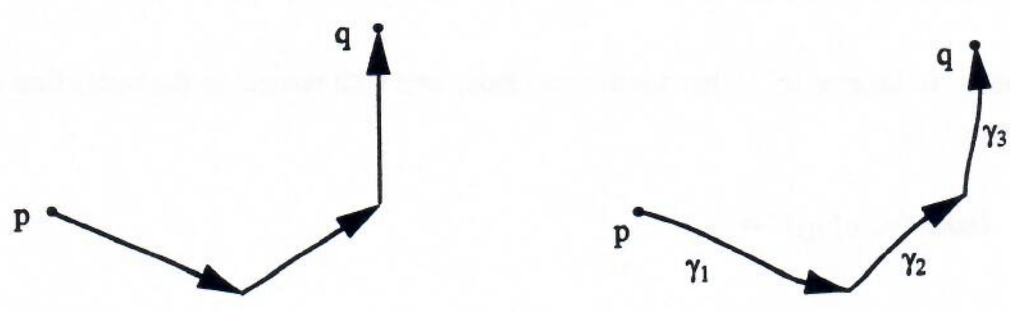

Figure 5. Two three legged paths.

$\gamma_3(0) = \gamma_2(t_2)$, and $\gamma_3(t_3) = q$? When $X_1 = \partial/\partial x$, $X_2 = \partial/\partial y$, and $X_3 = \partial/\partial z$, the answer is of course yes, and the neighborhood of q is covered smoothly by a multiflowbox.

Relevant here is Kneser's Theorem (~~19xx~~ 1923) about ODE's with non-unique solutions. It implies that the (forward) **solution funnel** for a continuous vector field X on $\mathbb{R}^m$,

$$F(p, s) = \{\gamma(t) : \gamma \text{ is an } X\text{-solution}, \gamma(0) = p, \text{ and } 0 \le t \le s\}$$

has a **cross-section**

$$K(p, s) = \{\gamma(s) : \gamma \text{ is an } X\text{-solution and } \gamma(0) = p\}$$

that is a **continuum** - i.e., a non-empty, compact, connected set. Correspondingly, for several vector fields we get a **multifunnel**. For instance, the funnel of two vector fields X_1, X_2, is

$$F(p, s_1, s_2) = \{q : \text{There exist an } X_1 \text{ solution } \gamma_1, \text{ an } X_2\text{-solution } \gamma_2, \text{ and times } t_1 \in [0, s_1],\ t_2 \in [0, s_2], \text{ such that } \gamma_1(0) = p,\ \gamma_2(0) = \gamma_1(t_1), \text{ and } \gamma_2(t_2) = q\}.$$

Its cross-section $K(p, s_1, s_2)$ results by setting $t_1 = s_1$ and $t_2 = s_2$. An easy generalization of Kneser's Theorem implies that this cross-section is a continuum. Also, the following continuity conditions are easy to check. As functions from $\mathbb{R}^m \times \mathbb{R}^2$ into $\mathcal{K}$, the space of continua in $\mathbb{R}^m$ equipped with the Hausdorff metric, the multifunnel and its cross-section,

$$(p, s_1, s_2) \mapsto F(p, s_1, s_2) \text{ and } (p, s_1, s_2) \mapsto K(p, s_1, s_2),$$

are upper semi-continuous. This means that a slight change of (p, s_1, s_2) can cause the continuum to become much smaller, but not much larger. Moreover, for each fixed (p, s_1) we have continuity in s_2. These properties generalize easily to three or more vector fields.

Question 2 amounts to asking whether the backwards X_3-funnel through $q = (1, 1, 1)$ meets the forward (X_1, X_2) - funnel through $p = (0, 0, 0)$. Using the upper semi-continuity

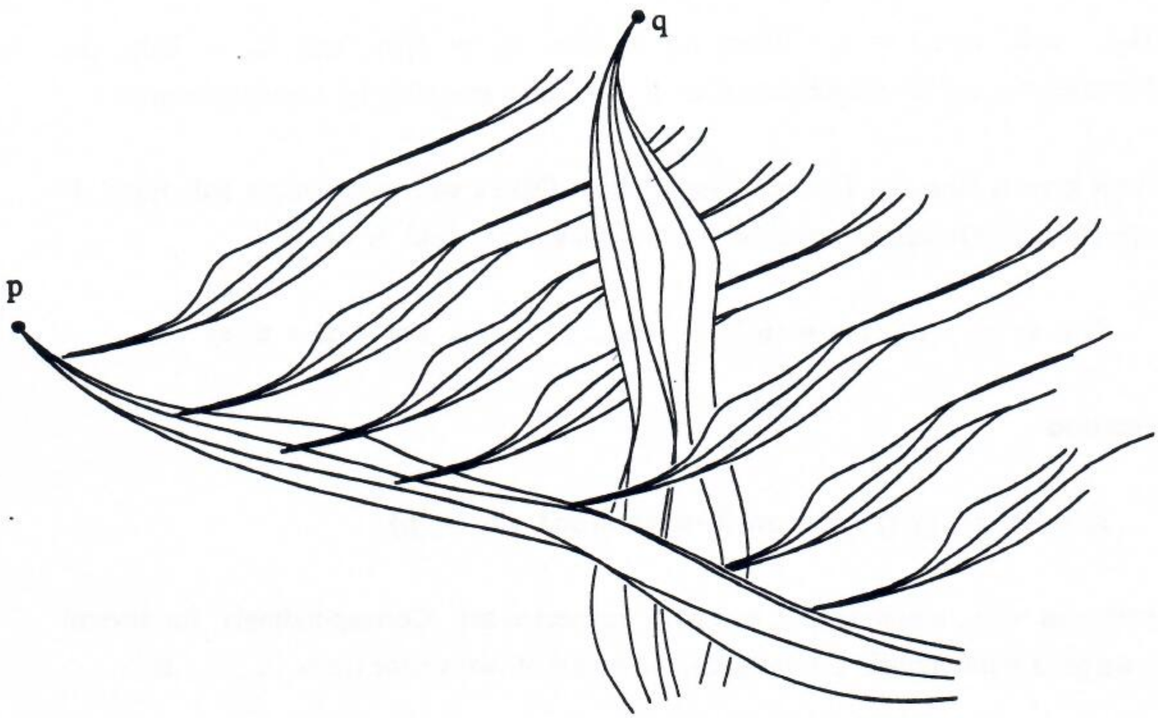

Figure 6. Must these funnels meet?

and continuum properties of the funnels, it is possible to give a geometric argument to show that the intersection does indeed exist, and this settles the test case.

We believe that the general case should also have an affirmative answer, and that it may involve some kind of "Čech intersection condition". Consider two smooth discs D_1 and D_2 in $\mathbb{R}^m$ that have dimensions d_1 and d_2. Assume that $d_1 + d_2 = m$ and D_1 intersects D_2 transversally at a point r. If D_1' and D_2' approximate D_1 and D_2 in an appropriate sense then they should intersect at a point r' that approximates r. When D_1' and D_2' are the images of D_1 and D_2 under continuous maps g_1 and g_2 that approximate the inclusion maps then it is not hard to see that the intersection does persist.

We are interested, however, in more general sets D_1' and D_2'. We want to let D_1' and D_2' be funnels, cross-sections of funnels, etc. As shown in Pugh (1975) funnel sections can be topologically pathological; for example they can fail to be locally connected. We hope that they are, or that they contain, "Čech discs" with nonzero "Čech intersection number".

References

D. V. Anosov. *Geodesic flows on closed Riemannina manifolds with negative curvature,* Proceedings of the Stecklov Institute #90, 1967.

R. Bowen. A horseshoe with positive measure, Invent. Math., 29 (1975), 203-204

K. A. Grasse. On Accessibility and Normal Accessibility: The Openness of Controllability in the fine C^0 topology, J. Diff. Eqs. **53** (1984), 387–414.

M. Grayson, C. Pugh, & M. Shub. GPS. Stably Ergodic Diffeomorphisms, Annals of Math. **140** (1994), 295–329.

M. Gromov. Carnot–Caratheodory Spaces Seen From Within, IHES preprint (1995).

M. Hirsch, C. Pugh, & M. Shub. HPS. *Invariant Manifolds,* Springer Lecture Notes in Math, #583, Heidelberg (1977).

H. Kneser. Über die Lösungen eines Systems gewöhnlichen Differentialgleichungen das der Lipschitzchen Bedingung nicht genucht, S-B Preuss Akad Wiss. Math Kl. (1923) 171-174.

A. Kolmogorov. Proceedings of the International Congress of Mathematicians 1 (1954), 315–313.

C. Lobry. Une propriété générique des couples de champs de vecteurs, Czechoslovak Math. J. **22** **(97)** (1972), 230–237.

C. Lobry. Dynamical Polysystems and Control Theory, in *Geometric Methods in Systems Theory*; Proceedings of the NATO Advanced Study Institute, London England, August 27 – September 7, 1973. (Eds. David Q. Mayne & R. W. Brockett), D. Reidel, Boston (1973).

C. Pugh. ~~Funnels~~ Sections. J. Diff. Eqs. 19 (1975), 270-295.

C. Pugh & M. Shub. Ergodicity of Anosov Actions, Inventiones Math. **15** (1972), 1–23.

C. Pugh, M. Shub, & A. Wilkinson. PSW. Hölder Foliations, to appear in Duke J. Math. (1996).

Pugh & Shub. Stably Ergodic dynamical systems and partial hyperbolicity, preprint (1996).

C. Robinson & L.–S. Young. Nonabsolutely continuous foliations for an Anosov diffeomorphism. Invent. Math 61 (1980), 159-176.

H. Sussman. Some properties of vector field systems that are not altered by small perturbations, J. Diff. Eqs. **20** (1976), 292–315.

A. Wilkinson. Stable Ergodicity of the Time One Map of a Geodesic Flow, thesis at the University of California at Berkeley, (1995).