On zeta functions and Anosov diffeomorphisms

by

Anthony Manning.

A thesis submitted for the Ph.D. degree at the University of Warwick, June 1972.

ABSTRACT.

This thesis considers some problems in Dynamical Systems concerned with zeta functions and with Anosov diffeomorphisms.

In chapter 1 Bowen's method of expressing a basic set of an Axiom A diffeomorphism as a quotient of a subshift of finite type is used to calculate the numbers of periodic points of the diffeomorphism and show that its zeta function is rational which gives an affirmative answer to a question of Smale.

The rest of the thesis is concerned with Anosov diffeomorphisms of nilmanifolds. Chapter 2 contains some facts about nilmanifolds describing them as twisted products of tori. A nilmanifold has a maximal torus factor. A hyperbolic nilmanifold automorphism projects onto an automorphism of this torus and we say it has the toral automorphism as a factor. In chapter 3 we generalize this situation to show that many diffeomorphisms of other manifolds have toral automorphisms as factors and give some examples.

In the last chapter we use a spectral sequence associated to another decomposition of a nilmanifold into tori to calculate the Lefschetz number of any diffeomorphism of the nilmanifold. This enables us to prove a necessary condition on the map induced by an Anosov diffeomorphism of a nilmanifold on its fundamental group. Then we consider the question of finding hyperbolic automorphisms of nilmanifolds

from the decomposition into tori. Finally we calculate the zeta function of such an automorphism.

CONTENTS.

		Page
	Abstract	
	Introduction	2
1.	Axiom A diffeomorphisms have rational zeta functions	8
1.1	Introduction	8
1.2	Markov partitions	9
1.3	The subshifts $\Lambda(A_i)$	12
1.4	The result	14
2.	Nilmanifolds	18
2.1	Preliminary facts	18
2.2	The torus decomposition using the lower central series	19
2.3	The torus decomposition using the upper central series	20
3.	Anosov diffeomorphisms as factors	23
3.1	Introduction	23
3.2	Finding the torus T^r and the quotient map k	24
3.3	Applications to Anosov diffeomorphisms	26
3.4	Applications to Axiom A diffeomorphisms	28
3.5	Examples	30
4.	Anosov diffeomorphisms on nilmanifolds	38
4.1	The induced map on the fundamental group	38
4.2	Summary of what is known about these diffeomorphisms	42
4.3	Hyperbolic automorphisms of nilmanifolds	43
4.4	Zeta functions	51
	References	53

INTRODUCTION.

The subject of differentiable dynamical sytems studies a diffeomorphism f or a flow φ_t on a smooth compact connected manifold M and in particular the global and geometric properties of the orbits of f or φ. The idea is to mimic a physical system with state space M and with $\varphi_t x$ the state of the system at time t if it is in state x at time 0.

It is sensible to study only flows φ whose orbits have the same properties as the orbits of every nearby flow ψ because a physical system is subject to small perturbations in the controlling forces and to inaccuracies of measurement so that we could not have been sure whether our system was in fact represented by φ or ψ. Such a flow φ is said to be stable. For some years the main problem in this subject was to find a precise definition of stability so that stable flows were dense in the C^1-topology in the space of all flows on M and also that stable flows were amenable to some form of classification. It is best to tackle problems on flows for diffeomorphisms first because it is easier to work with diffeomorphisms and yet not too hard to extend results proved for diffeomorphisms to flows - see the remark on this at the end of chapter 1.

The problem is then to investigate the orbit structure of a diffeomorphism f having some stability property. One easily noticed form of behaviour of our system will be regularly repeated behaviour. This corresponds to a periodic point of f i.e. a point x s.t. $f^n x=x$ for some n. An important question about f will be to discover how many periodic points of each period n it has and ensure this is the same for all diffeomorphisms close enough to f.

Define a closed subset A of M to be an <u>attractor</u> for f if it is contained in some open set U and $A = \bigcap_{n\geq 0} f^n U$. That is all points near A approach A under repeated application of f. An attractor of f will be of particular interest because it determines the behaviour of the system started anywhere within an open subset of the state space, in fact anywhere within its basin of attraction $\bigcup_{n\in\mathbb{Z}} f^n U$.

Under this definition the whole manifold M is an attractor for any f whereas we really want to investigate the simplest pieces into which the attractors of f can be broken down. So we impose the condition of topological transitivity on an attractor. $f|A$ is <u>topologically transitive</u> if $\exists x\in A$ s.t. $A = \mathrm{cl}\{f^n x; n\in\mathbb{Z}\}$. Then A acts as a whole and we should like to understand its structure. There will be only a finite number of topologically transitive attractors if we impose the condition that f satisfy Axiom A [29]. This was Smale's candidate for a condition on diffeomorphisms that would ensure they had a particular kind of stability called Ω-stability and were amenable to classification. Unfortunately in [31] he found that these diffeomorphisms were not dense in the space of all diffeomorphisms but Franks [10] and Guckenheimer [12] have found slightly stronger stability conditions that imply Axiom A.

Before proceeding with this discussion of dynamical systems we had better give the main definitions. Let M be a smooth compact connected manifold without boundary and Diff(M) the space of C^1 diffeomorphisms of M with the C^1-topology. Let $f \in \mathrm{Diff}(M)$.

Definition. Let $N_m(f)$ be the number of fixed points of f^m. The zeta function of f is

$$\zeta(f,t) = \exp \sum_{m=1}^{\infty} (1/m) N_m(f) t^m.$$

Definition. The nonwandering set $\Omega(f) = \{x \in M;$ for every neighbourhood U of x $\exists n$ s.t. $f^n U \cap U \neq \emptyset\}$.

Definition. f satisfies Axiom A if

(a) the restriction of the tangent bundle to Ω $T_\Omega M$ has two continuous Df-invariant subbundles E^s, E^u with $T_\Omega M = E^s \oplus E^u$ and for any Riemannian metric $\exists$ constants c, λ $c>0$, $0<\lambda<1$ s.t.

$$\|Df_x^n | E_x^s\| < c\lambda^n \quad \text{and}$$
$$\|Df_x^{-n} | E_x^u\| < c\lambda^n \quad \forall x \in \Omega, n \geq 0$$

and (b) the periodic points of f are dense in Ω.

Theorem. (Smale's Spectral decomposition theorem) If f satisfies Axiom A then Ω can be written as $\Omega_1 \cup \ldots \cup \Omega_c$ where the Ω_i are closed disjoint f-invariant subsets on each of which f is topologically transitive. The sets Ω_i are called basic sets.

Definition. f is an Anosov diffeomorphism if it satisfies Axiom A (a) with Ω replaced by M.

Definition. $f, g \in \mathrm{Diff}(M)$ are topologically conjugate if $\exists$ a homeomorphism $h: M \to M$ s.t. $hf = gh$.

Definition. f is structurally stable if it has a neighbourhood in Diff(M) consisting of diffeomorphisms topologically conjugate to it.

Theorem. (Anosov [1]) Any Anosov diffeomorphism is structurally stable.

However, few examples of Anosov diffeomorphisms are known and these only on tori and nilmanifolds (and manifolds finitely covered by them) - the so-called hyperbolic toral automorphisms and hyperbolic nilmanifold automorphisms, see e.g. [8]. Franks [8] and Newhouse [21] have shown that if an Anosov diffeomorphism f has E^s or E^u 1-dimensional then f is a hyperbolic toral automorphism but there are not many other results about Anosov diffeomorphisms of an arbitrary manifold.

Part of the attractiveness of the subject of differentiable dynamical systems lies in the fact that a wide range of tools from other branches of mathematics can be used to attack its problems. Moser's proof of the structural stability of Anosov diffeomorphisms as expounded by Mather [29; pp.792-5] uses a manifold of maps and an implicit function theorem. Proofs of the existence of stable manifolds have used the stability properties of hyperbolic automorphisms of Banach spaces, see e.g. [15]. Sinai

[28] and Bowen [3,4 and 5] have looked at connections with measure theory, entropy and topological dynamics. Smale [29] used the idea of Lefschetz number from algebraic topology to count periodic points. In [30] he used a handle decomposition of M to show that any f is isotopic to one satisfying Axiom A. And Franks [8] approached Anosov diffeomorphisms via their homotopy theoretic properties.

In this thesis chapter 1 uses Bowen's work to calculate $\zeta(g,t)$ for an Axiom A diffeomorphism g and show it is a rational function of t. Chapter 2 contains some facts about nilmanifolds that are needed later, describing in particular how a nilmanifold decomposes into tori. Chapter 3 uses work of Franks to show that if f induces a hyperbolic map on $H^1(M;\mathbb{Z})$ then $f:M\to M$ has a hyperbolic toral automorphism as a factor and then gives some examples. In chapter 4 M is a nilmanifold and $f:M\to M$ is an Anosov diffeomorphism. §4.2 summarises what is known about such f after a theorem in §4.1 which uses Lefschetz numbers and a spectral sequence to obtain a necessary condition on $f_*:\pi_1(M)\to\pi_1(M)$. §4.3 considers hyperbolic automorphisms of nilmanifolds from the algebraic point of view and §4.4 calculates the zeta function of such an automorphism.

I should like to record my grateful thanks to my wife Hanne for all her support and encouragement during the last three years, to the two supervisors I have had in the course of this work, Professor R. Bowen and Professor E. C. Zeeman, for introducing me to dynamical systems and for much guidance and help and to the Science Research Council who supported me financially.

Chapter 1. AXIOM A DIFFEOMORPHISMS HAVE RATIONAL ZETA FUNCTIONS.

1.1 Introduction.

This result was conjectured by Smale [29; p.785] and its value is to show that the numbers of periodic points of all orders of the diffeomorphism are determined by the finite number of zeros and poles of the zeta function. In [34] Williams gives a survey of results on this function up to 1968. Then in [11] Guckenheimer showed by using a double cover and the Lefschetz Trace Formula that an Axiom A diffeomorphism has rational zeta function provided it satisfies the no cycle property. Since the theorem of this chapter was first proved Simon [27] has found a set (with non-empty interior) of diffeomorphisms not satisfying Axiom A whose zeta functions are not rational.

In [4] Bowen, following Sinai [28], proved the existence of a Markov partition of a basic set Ω_b of an Axiom A diffeomorphism by means of which Ω_b can be expressed as a quotient of a subshift of finite type. Since the existence of Markov partitions does not depend on the no cycle property and the zeta function of a subshift of finite type is known from [7] this seems a natural method for approaching the zeta function. In this chapter this partition is used in §3 to construct new subshifts by means of which the periodic points of Ω_b can be counted in §4. This chapter has appeared in the Bulletin of the London Mathematical Society volume 3 (1971) pages 215-220.

1.2 Markov Partitions.

Let $g:M \to M$ be an Axiom A diffeomorphism and recall from the introduction Smale's Spectral Decomposition Theorem which says that $\Omega(g) = \Omega_1 \cup \ldots \cup \Omega_c$ where each basic set Ω_b is closed and g-invariant and $g|\Omega_b$ is topologically transitive. It is clear that $N_m(g) = \Sigma_{b=1}^{c} N_m(g|\Omega_b)$ so that, as in [29; p.766],

$$\begin{aligned} \zeta(g,t) &= \exp \Sigma_{m=1}^{\infty}(1/m)\left\{\Sigma_{b=1}^{c} N_m(g|\Omega_b)\right\}t^m \\ &= \exp \Sigma_{b=1}^{c} \Sigma_{m=1}^{\infty}(1/m)N_m(g|\Omega_b)t^m \\ &= \Pi_{b=1}^{c} \zeta(g|\Omega_b,t) . \end{aligned} \quad (1)$$

Thus it is sufficient to prove that $\zeta(g|\Omega_b,t)$ is rational for each b.

Let $f = g|\Omega_b$ for some fixed b. Then $f:\Omega_b \to \Omega_b$ is expansive with expansive constant $\varepsilon > 0$ say. This means that, for any two distinct points $x,y \in \Omega_b$, there is n s.t. $d(f^n x, f^n y) > \varepsilon$. Then according to [4] there is a Markov partition for Ω_b, that is a finite cover $\mathcal{E}$ of Ω_b by closed subsets called rectangles whose diameters we require here to be less than $\varepsilon/2$. The rectangles are pairwise disjoint except possibly for the intersection of their boundaries. And if $E_j \in \mathcal{E}$ then (figure 1.1)

$$x,y \in E_j \Rightarrow W^s(x,\varepsilon) \cap W^u(y,\varepsilon) \in E_j .$$

Finally the rectangles satisfy the 'Markov property'

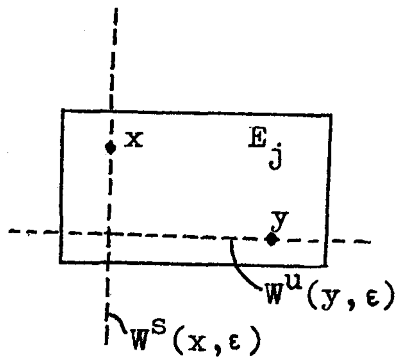

Figure 1.1

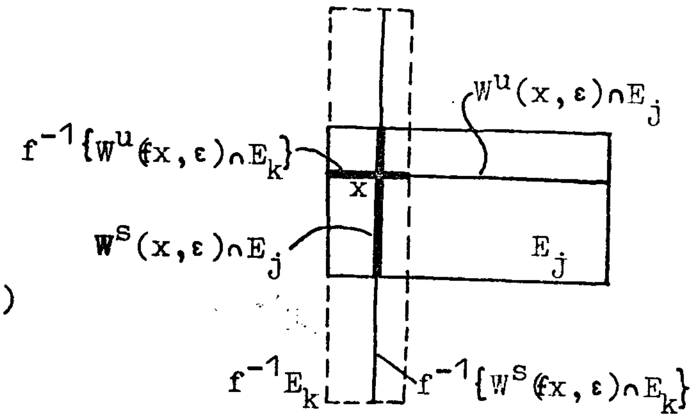

Figure 1.2

that if $E_j, E_k \in \mathcal{C}$ and $x \in \mathrm{int}E_j \cap f^{-1}(\mathrm{int}E_k)$ then (figure 1.2)

and
$$\left.\begin{aligned} W^s(x,\varepsilon)\cap E_j &\subset f^{-1}\{W^s(fx,\varepsilon)\cap E_k\} \\ W^u(x,\varepsilon)\cap E_j &\supset f^{-1}\{W^u(fx,\varepsilon)\cap E_k\} \end{aligned}\right\} \quad (2)$$

Define the transition matrix $T = (t(E_j,E_k))$ by $t(E_j,E_k) = 1$ if
$$f(\mathrm{int}E_j)\cap \mathrm{int}E_k \neq \emptyset$$
and $t(E_j,E_k) = 0$ otherwise. Then T gives rise to a subshift of finite type $\tau:\Lambda(T)\to\Lambda(T)$ as follows. Let $E = (E_n)_{n=-\infty}^{\infty}$ be a sequence of elements of $\mathcal{C}$ such that $t(E_n,E_{n+1}) = 1$ for all n. Let $\Lambda(T)$ = the set of all such sequences. Let $\tau:\Lambda(T)\to\Lambda(T)$ be defined by $\tau E = F$ where $F_n = E_{n+1}$ for all n. Now the map $\pi:\Lambda(T)\to\Omega_b$ given by $\pi E = \bigcap_{n=-\infty}^{\infty} f^{-n}E_n$ is well defined by the conditions (2) and expansiveness. π is one to one almost everywhere and gives a commutative diagram

$$\begin{array}{ccc} \Lambda(T) & \xrightarrow{\tau} & \Lambda(T) \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ \Omega_b & \xrightarrow[f]{} & \Omega_b \end{array} .$$

The interested reader may consult [4] for proofs and more details of these Markov partitions.

From a knowledge of this subshift and of which rectangles intersect it is possible to calculate $\zeta(f,t)$. The rectangles $E_1,\dots,E_r$ are said to be <u>related</u> if $E_1 \cap \dots \cap E_r \neq \emptyset$. The following lemma was shown me by Bowen.

<u>Lemma 1</u>. Let $E^1,\dots,E^r$ be distinct elements of $\Lambda(T)$. If, for each n, the rectangles $E_n^1,\dots,E_n^r$ are related then $\pi E^1 = \pi E^2 = \dots = \pi E^r$.

<u>Proof</u>. Let $\pi E^j = x^j \in \Omega_b$. $f^n x^j \in E_n^j$, $f^n x^k \in E_n^k$. So

$$d(f^n x^j, f^n x^k) \leqslant \operatorname{diam}(E_n^j \cup E_n^k) < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

for all n. But ε is an expansive constant for f. So $x^j = x^k$ as required.

In [5] by a detailed investigation for arbitrary $x \in \Omega_b$ of the T-relationship between the rectangles to which x belongs and those to which fx belongs Bowen shows

<u>Proposition 10</u>. There is an integer d such that, for all x, $\pi^{-1}x$ has at most d elements.

<u>Corollary 11</u>. E is a periodic element of $\Lambda(T)$ if and only if πE is a periodic point of Ω_b under f.

<u>Proposition 12</u>. If $\pi E = \pi F$ is periodic under f and if $E_n = F_n$ for some n then $E_n = F_n$ for all n.

These results are needed for

<u>Lemma 2</u>. If $E^1,\dots,E^r$ are as in Lemma 1 and each is

periodic under τ (with possibly different periods) then $t(E_n^j, E_{n+1}^k) = 0$ for $j \neq k$.

<u>Proof</u>. By Proposition 12 of [5] for each n $E_n^1, \ldots, E_n^r$ are distinct. Choose $m>0$ such that $\tau^m E^j = E^j$ for each j. Suppose contrary to the Lemma that there are j,k,h such that $j \neq k$ and $t(E_h^j, E_{h+1}^k) = 1$. Figure 1.3 shows part of the directed graph with vertex set $\mathcal{E}$ and adjacency matrix T. Related rectangles are joined by dotted lines. A point of $\Lambda(T)$ is a two-way infinite path in this graph. E^j, E^k are paths round the inner and outer squares respectively. (m = 4 in the figure.)

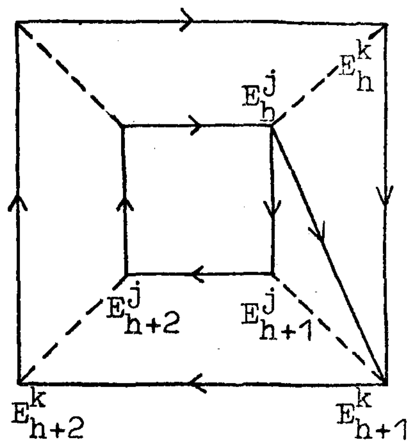

Figure 1.3

For each integer p construct points v^p of $\Lambda(T)$ with paths going round the inner square until time h+pm and then round the outer square. More precisely, $v_n^p = E_n^j$ for $n \leqslant h+pm$, $v_n^p = E_n^k$ for $n > h+pm$. Then all the points v^p are distinct and by Lemma 1 they have the same image under π. But this contradicts Proposition 10 of [5] and so $t(E_h^j, E_{h+1}^k) = 0$.

1.3 <u>The Subshifts $\Lambda(A_i)$</u>.

$\pi\tau = f\pi$ implies that $\pi\{\mathrm{Fix}(\tau^m)\} \subset \mathrm{Fix}(f^m)$ but $N_m(\tau) \neq N_m(f)$ for two reasons:

(1) At the boundaries of the rectangles, that is where they intersect, π is many to one so several

points of $\mathrm{Fix}(\tau^m)$ may be mapped to the same point of $\mathrm{Fix}(f^m)$.

(2) If x has period m under f then f^m may rotate or reflect the manifold in the neighbourhood of x. If, further, $x \in E_j \cap E_k$ then it might for example interchange E_j and E_k and consequently τ^m would interchange the elements of $\pi^{-1}x$ containing E_j and E_k. Therefore in this case these elements of $\pi^{-1}x$ would have period 2m rather than m.

To capture the points x that have several inverse images the obvious thing to do is to construct subshifts whose symbols are sets of r related rectangles for various r. Unfortunately this approach is too simple because we find ourselves counting the points x too often, and so we need an algebraic device of k-tuples of sets of related rectangles to cancel out the overcounting. For the moment we confine ourselves to the formal definitions and the reasons will become apparent in the next section.

Define q to be the largest integer such that there is a set of q related rectangles. Fix k between 1 and q. Let $i = (i_1, \ldots, i_k)$ be a fixed k-tuple of positive integers and put $|i| = \sum i_j$. We suppose that $|i| \leq q$. Now for each $j = 1, \ldots, k$ let e_j be a set of i_j related rectangles. Let $u = (e_1, \ldots, e_k)$ be a k-tuple of such sets, such that the rectangles in $\bigcup e_j$ are all distinct and all related. Let $\mathcal{A}_i$ be the set of all such k-tuples u.

We now proceed to define a transition matrix A_i for the symbol set $\mathcal{A}_i$, induced by the original transition matrix T. Let $e_j = \{E^1,\ldots,E^p\}$, $f_j = \{F^1,\ldots,F^p\}$ be two sets of p related rectangles. Write $t(e_j,f_j)=1$ if there is a relabelling of the F's such that $t(E^h,F^h)=1$ for $h=1,\ldots,p$. Note that by Lemma 2 any such relabelling must be unique. Write $t(e_j,f_j)=0$ otherwise. Now given two elements of $\mathcal{A}_i$ $u=(e_1,\ldots,e_k)$, $v=(f_1,\ldots,f_k)$ write $a(u,v)=1$ if $t(e_j,f_j)=1$ for $j=1,\ldots,k$ and write $a(u,v)=0$ otherwise. This defines the transition matrix A_i for $\mathcal{A}_i$. From the symbol set $\mathcal{A}_i$ and the transition matrix A_i construct in the usual way the subshift of finite type $\alpha_i:\Lambda(A_i)\to\Lambda(A_i)$.

Remark. $\mathcal{C}=\mathcal{A}_1$, $T=A_1$ and $\tau:\Lambda(T)\to\Lambda(T)$ is the same as $\alpha_1:\Lambda(A_1)\to\Lambda(A_1)$. Moreover in the case where each $i_j=1$ A_i may be obtained as that submatrix of the tensor product of T with itself k times corresponding to those rows and columns that belong to $\mathcal{A}_i$.

1.4 The Result.

Theorem 1. $N_m(f)=\sum(-1)^{k+1}N_m(\alpha_i)$ where the sum is taken over all $i=(i_1,\ldots,i_k)$ with $1\leqslant k\leqslant q$ and $|i|\leqslant q$.

Proof. Any point of $\Lambda(A_i)$ gives rise in a natural way to $|i|$ points of $\Lambda(T)$ as follows. Let $y=(y_n)_{n=-\infty}^{\infty}$ be a point of $\Lambda(A_i)$ and let $y_0=(e_1,\ldots,e_k)$. Then each $E^h\in\bigcup e_j$ determines a unique point $z^h\in\Lambda(T)$; for instance $z_0^h=E^h$, $z_1^h=$the unique $F^h\in\bigcup f_j$, where $y_1=(f_1,\ldots,f_k)$, such that $t(E^h,F^h)=1$, and z_2^h, $z_3^h,\ldots$, z_{-1}^h, $z_{-2}^h,\ldots$ are defined inductively. Therefore z^h is uniquely determined, $1 \leq h \leq |i|$.

Now $\pi z^h = z \in \Omega_b$ where z is independent of h by Lemma 1. Define $\varphi : \Lambda(A_i) \to \Omega_b$ by $\varphi y = z$. Then $\varphi \alpha_i = f\varphi$. If $y \in \mathrm{Fix}(\alpha_i^m)$ then $\varphi y \in \mathrm{Fix}(f^m)$. So every point counted in $N_m(\alpha_i)$ corresponds under φ to a point counted in $N_m(f)$.

So it is sufficient to show that, for each $z \in \mathrm{Fix}(f^m)$, $\Sigma(-1)^{k+1} N'_m(\alpha_i) = 1$ where N'_m counts only points in $\varphi^{-1}z$. Let $\pi^{-1}z = \{z^1, \ldots, z^r\}$. If $|i| > r$ then $\varphi^{-1}z \cap \mathrm{Fix}(\alpha_i^m) = \emptyset$. The remaining i may be divided into three sets B, C, D thus

$$B = \{i ; |i| < r\}$$

$$C = \{i ; |i| = r \text{ and } k > 1\}$$

$$D = \{i ; k=1,\ i_1 = r\} = \{r\}.$$

Define $\psi : B \to C$ by $\psi(i_1, \ldots, i_k) = (i_1, \ldots, i_k, r - i_1 - \ldots - i_k)$. ψ is a bijection. ψ can be used to show that

$$\Sigma_B (-1)^{k+1} N'_m(\alpha_i) + \Sigma_C (-1)^{k+1} N'_m(\alpha_i) = 0$$

as follows.

Let x be a point of $\mathrm{Fix}(\alpha_i^m)$ with $\varphi x = z$ and $i \in B$. Define y in $\mathrm{Fix}(\alpha_{\psi i}^m)$ with $\varphi y = z$ thus: the first k sets of y_n are the k sets of x_n in that order and the $(k+1)$-th set of y_n is the set of $r - i_1 - \ldots - i_k$ rectangles in $\{z_n^1, \ldots, z_n^r\}$ but not in a set of x_n. By Lemma 2, $t(z_n^h, z_{n+1}^h) = 1$ but $t(z_n^h, z_{n+1}^p) = 0$ if $h \neq p$. The rectangles in x_n are paired with those in x_{n+1} according to the matrix T so those in the $(k+1)$-th set of y_n can be paired with those in the $(k+1)$-th set of y_{n+1}. Thus $a(y_n, y_{n+1}) = 1$ and $y \in \Lambda(A_{\psi i})$. Moreover, for each n, $x_n = x_{n+m}$ and $f^m z = z$ so that

$$\{z_n^1,\ldots,z_n^r\} = \{z_{n+m}^1,\ldots,z_{n+m}^r\},$$

so $y_n = y_{n+m}$ and $y \in \mathrm{Fix}(\alpha_{\psi i}^m)$.

Similarly given $y \in \mathrm{Fix}(\alpha_{\psi i}^m) \cap \varphi^{-1}z$ obtain $x \in \mathrm{Fix}(\alpha_i^m) \cap \varphi^{-1}z$ by simply omitting the (k+1)-th set of y_n to get x_n. These two operations are mutually inverse and give a bijection $\mathrm{Fix}(\alpha_i^m)\cap\varphi^{-1}z \to \mathrm{Fix}(\alpha_{\psi i}^m)\cap\varphi^{-1}z$. Therefore

$$(-1)^{k+1}N_m'(\alpha_i) + (-1)^{k+2}N_m'(\alpha_{\psi i}) = 0 \quad \text{and} \quad \Sigma_B + \Sigma_C = 0.$$

It remains to prove that $\Sigma_D(-1)^{k+1}N_m'(\alpha_i) = 1$. The only point in $\varphi^{-1}z \cap \Lambda(A_r)$ is w where $w_n = \{z_n^1,\ldots,z_n^r\} = \{E \in \mathcal{C};\, f^n z \in E\}$. $w_{n+m} = \{E \in \mathcal{C};\, f^{n+m} z \in E\} = \{E \in \mathcal{C};\, f^n z \in E\} = w_n$, so $w \in \mathrm{Fix}(\alpha_r^m) \cap \varphi^{-1}z$ and

$$\Sigma_D(-1)^{k+1}N_m'(\alpha_i) = (-1)^2 N_m'(\alpha_r) = 1$$

as required. This concludes the proof.

<u>Corollary</u>. $\zeta(f) = \prod_{|i| \le q} \zeta(\alpha_i)^{(-1)^{k+1}}$.

<u>Proof</u>. This follows from theorem 1 by the argument of (1) above.

<u>Theorem 2</u>. $\zeta(g)$ is rational if g satisfies Axiom A.

<u>Proof</u>. From [7] $\zeta(\alpha_i)$ is rational. In fact $\zeta(\alpha_i) = \{\det(I - tA_i)\}^{-1}$. Hence, from the Corollary, $\zeta(f)$ is rational, and $\zeta(g)$ is just the product of c functions like $\zeta(f)$.

<u>Question</u>. The zeta function of a toral Anosov diffeomorphism with eigenvalues $\lambda_1,\ldots,\lambda_n$ is given in [29; p.769] as a product and quotient of terms $(1-\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_k}t)$. The same formula applies to the

nilmanifold examples (see §4.4). If we are given the zeta function as $\prod_i \zeta(\alpha_i)^{(-1)^{k+1}}$ is it possible to recover the original eigenvalues $\lambda_1,\ldots,\lambda_n$ from those of the matrices A_i?

Remark. The simplest basic sets for Axiom A diffeomorphisms are the 0-dimensional ones which are just subshifts of finite type. As described in §1.2 any other basic set can be expressed as a quotient of a subshift of finite type. Recently Bowen has been extending his work to Axiom A flows. Here he finds that the simplest type of basic set (apart from fixed points) is 1-dimensional (actually the suspension of a subshift of finite type at a time which varies in a Lipschitz manner). Any other basic set is a finite-to-one quotient of one of these. In [6] he makes use of the methods of this chapter, in particular theorem 1, to show that the zeta function of an Axiom A flow is a product and quotient of zeta functions of certain 1-dimensional basic sets. This and Bowen's other successes with Axiom A flows illustrate the process of obtaining results for diffeomorphisms and then extending them to flows.

Chapter 2. NILMANIFOLDS.

2.1 Preliminary Facts.

This chapter contains some facts about nilmanifolds which will be needed later on.

Definition. A nilmanifold is a compact homogeneous space N/D where N is a connected simply connected nilpotent Lie group and D is a uniform discrete subgroup of N.

Malcev [19] investigated nilmanifolds in some detail and we quote the following two results from his paper.

Fact 1. Nilmanifolds are determined by their fundamental group D. For an abstract group D to be the fundamental group of some nilmanifold it is necessary and sufficient that D be finitely generated torsion-free and nilpotent.

Fact 2. If N/D is a nilmanifold then any automorphism of the group D can be uniquely extended to an automorphism of N.

This automorphism preserves the subgroup D and induces a diffeomorphism of N/D which we may call a nilmanifold automorphism.

N/D is a K(D,1) since its universal cover N is contractible, see [25; p.180]. From theorem 8.1.11 of [32] we get immediately

Fact 3. $f_1, f_2 : N/D \to N/D$ are freely homotopic if and only if the endomorphisms they induce on the fundamental group D are conjugate i.e. differ by an inner automorphism of D.

Our interest in nilmanifolds arises from the fact that some of them admit Anosov diffeomorphisms - the so-called hyperbolic nilmanifold automorphisms. The simplest manifolds known to admit Anosov diffeomorphisms are the tori and a nilmanifold can be expressed as a 'twisted product' of tori. There are two methods of decomposing a nilmanifold into tori.

2.2 The Torus Decomposition Using the Lower Central Series.

The first method is described by Parry in [22]. We recall only as much as we shall need in the next chapter. Let $N^1=[N,N]$=the subgroup of N generated by elements of the form $x^{-1}y^{-1}xy$ for any $x,y\in N$. There is an obvious projection from N/D to the space $N/N^1\cdot D$. (The dot denotes semidirect product.) The space $N/N^1\cdot D$ is a torus isomorphic to $(N/N^1)/(N^1\cdot D/N^1)$. Its universal covering space is N/N^1 and its fundamental group is $N^1\cdot D/N^1$ which is isomorphic to $D/(N^1\cap D)$.

Let us investigate the group $D/(N^1\cap D)$ more closely. $D^1=[D,D]$ is clearly a subgroup of $N^1\cap D$ and by Malcev's description of D as "spanning" N we see that D^1 can only have finitely many cosets in $N^1\cap D$. Now D/D^1 is an abelian group and $D/(N^1\cap D)$ is a free abelian group. We deduce that $D/(N^1\cap D)$, the fundamental group of the torus $N/N^1\cdot D$, is the quotient of D/D^1 by its torsion subgroup. Notice that $D=\pi_1(N/D)$ so $D/D^1=H_1(N/D;\mathbf{Z})$. The torus $N/N^1\cdot D$ is known as the <u>maximal torus factor</u> of the nilmanifold N/D. We use the universal coefficient theorem (e.g. 5.4.13c of [13]) by which

$$H^1(M;\mathbb{Z}) = H_1(M;Z)\pitchfork\mathbb{Z}$$

and recall that the group of homomorphisms from the torsion subgroup of H_1 to $\mathbb{Z}$ is trivial. Thus

$$H^1(N/D;\mathbb{Z}) \cong D/D^1\pitchfork\mathbb{Z} \cong D/(N^1\cap D)\pitchfork\mathbb{Z} \cong H_1(N/N^1\cdot D;\mathbb{Z})\pitchfork\mathbb{Z} \cong H^1(N/N^1\cdot D;\mathbb{Z}).$$

We sum this up as

Proposition 1. A nilmanifold N/D has a maximal torus factor $N/N^1\cdot D$ and $H^1(N/D;\mathbb{Z}) = H^1(N/N^1\cdot D;\mathbb{Z})$.

2.3 The Torus Decomposition Using the Upper Central Series.

I should like to thank Professor W. Parry for discussions on this section. In chapter 4 we shall need this second method of decomposing a nilmanifold into tori which goes as follows.

Definition. The upper central series

$$\{e\} = G_0 \subset G_1 \subset G_2 \subset \ldots \subset G$$

of a group G is defined inductively. G_1 is the centre of G. Let p_i be the projection of G onto G/G_i. G_{i+1} is defined to be p_i^{-1}(the centre of G/G_i).

For the nilpotent groups N and D the upper central series $\{e\} = N_0 \subset N_1 \subset \ldots \subset N_c = N$ and $\{e\} = D_0 \subset D_1 \subset \ldots \subset D_c = D$ have finite length c. (That the two series have the same length could be proved as a corollary to Lemma 1.) We shall use these central series to find a torus which acts on N/D with quotient space another nilmanifold. The torus will be N_1/D_1. Let $\mathfrak{n}$ be the Lie algebra of N and $\exp: \mathfrak{n} \to N$ the exponential map (which is injective).

Lemma 1. $N_1 \cap D = D_1$.

Proof. Clearly $N_1 \cap D \subset D_1$ and $D_1 \subset D$. We must show that $D_1 \subset N_1$ or $\exp^{-1} D_1 \subset \exp^{-1} N_1 =$ the centre of $\mathfrak{n}$. But if $\xi \in \exp^{-1} D_1$ then $[\xi, \eta] = 0$ for all $\eta \in \exp^{-1} D$, and we know that $\exp^{-1} D$ spans $\mathfrak{n}$ as a vector space. Thus $[\xi, \eta] = 0$ for all $\eta \in \mathfrak{n}$ as required.

Lemma 2. $N_1 / N_1 \cap D \cong N_1 \cdot D / D$.

Proof. The obvious map $x(N_1 \cap D) \mapsto xeD$ for $x \in N_1$ is an isomorphism.

Lemma 3. N_1 / D_1 is compact and a torus.

Proof. Let P be the vector space over the rationals spanned by $\exp^{-1} D$. Then P with the bracket of $\mathfrak{n}$ is a rational Lie algebra (the one discussed by Malcev in §4 of [19]) and $\mathfrak{n} = P \otimes R$. Clearly $\exp^{-1} D_1$ spans P_1, the centre of P. Let $\delta_1, \ldots, \delta_n \in \exp^{-1} D$ span P. Then the centre of $\mathfrak{n}$ is $\mathfrak{n}_1 = \bigcap_{i=1}^{n} \ker(\mathrm{ad}_{\delta_i} : \mathfrak{n} \to \mathfrak{n})$ and, see [16; p.28]; this is

$$\bigcap_{i=1}^{n} \{\ker(\mathrm{ad}_{\delta_i} : P \to P) \otimes R\} = \{\bigcap_{i=1}^{n} \ker(\mathrm{ad}_{\delta_i} : P \to P)\} \otimes R = P_1 \otimes R.$$

So $\mathfrak{n}_1 = P_1 \otimes R$ and N_1 / D_1 is compact. N_1 / D_1 is a torus because it is the compact quotient space of a Euclidean space N_1 by a free abelian group D_1.

Lemma 4. $N_1 \cdot D$ is closed in N.

Proof. Take sequences $x_n \in N_1$, $d_n \in D$ s.t. $x_n d_n \to m \in N$ as $n \to \infty$. Choose $e_n \in D_1$ s.t. $x_n e_n$ is contained in a compact set which is possible by Lemma 3. Then, taking a subsequence if necessary, $x_n e_n \to x$, say, in N_1. But $x_n e_n e_n^{-1} d_n \to m$ so $e_n^{-1} d_n$ is a convergent sequence in the discrete group D. $\therefore e_n^{-1}d_n=d$, say, in D for $n>$some n_0. $xd=m$ so $m\in N_1\cdot D$ as required.

Now the torus N_1/D_1 acts on our nilmanifold N/D by $(xD_1,yD)\mapsto xyD$ for $x\in N_1, y\in N$. This is well-defined since elements of D_1 commute with y. The orbit space of this action is $N/N_1\cdot D$ which is $(N/N_1)/(D/D_1)$ another nilmanifold. (By Lemma 4 $N/N_1\cdot D$ has the quotient topology.) When we use the same method to construct a torus acting on this nilmanifold we see from the definition of the upper central series that it is precisely $(N_2/N_1)/(D_2/D_1)$ and the orbit space is the nilmanifold $(N/N_2)/(D/D_2)$. Repeating this procedure we get

Proposition 2. The nilmanifold N/D is the extension of N_1/D_1 by $(N_2/N_1)/(D_2/D_1)$ by ... by $(N/N_{c-1})/(D/D_{c-1})$. Each of these spaces is a torus and their fundamental groups are the quotient groups D_i/D_{i-1} $i=1,2,\ldots,c$ of the upper central series of D.

Recall theorem 2 of [20] which says that D_i/D_{i-1} is free abelian for each i, a fact which is not necessarily true of the lower central series of D. The nilmanifold N/D may be called a c-step nilmanifold.

Chapter 3. ANOSOV DIFFEOMORPHISMS AS FACTORS.

3.1 Introduction.

In this chapter we shall consider some commuting diagrams

$$\begin{array}{ccc} M & \xrightarrow{f} & M \\ k\downarrow & & \downarrow k \\ T^r & \xrightarrow[g]{} & T^r \end{array}$$

where g is a hyperbolic toral automorphism, f is a diffeomorphism of M and $k:M \to T^r$ is a continuous map. We shall say that g is a factor of f borrowing the word from the measure theorists' terminology.

The motivation for this chapter comes from

Problem. Given an Anosov diffeomorphism $f:M \to M$ is f topologically conjugate to a hyperbolic nilmanifold automorphism?

If $f^*:H^1(M;\mathbb{Z}) \to H^1(M;\mathbb{Z})$ is hyperbolic (i.e. any element of $GL(r,\mathbb{Z})$ representing it has no eigenvalues of modulus 1) we shall set up a commutative diagram of the above form in which, if M is a nilmanifold, the torus T^r must be the maximal torus factor of M described in chapter 2. Three more stages would be necessary to solve the problem above but we have not made any progress with them.

(1) Show that f^* must be hyperbolic. In this direction Hirsch [14] has shown that f^* cannot have a root of unity as an eigenvalue under certain conditions

on M (every infinite cyclic cover of M must have finite dimensional rational homology).

(2) Show that k is the projection of a continuous fibre bundle and that the restriction of f to the fibre is again an Anosov diffeomorphism of a manifold so that we can apply the procedure again. Eventually the fibre would have dimension ≤ 3 and it is known [8 and 21] that this would have to be a hyperbolic toral automorphism. In this way M would be expressed as a sequence of torus extensions and these extensions could be examined to check that M is a nilmanifold.

(3) Modify this procedure to take account of the fact that M might be an infranilmanifold.

3.2 Finding the Torus T^r and the Quotient Map k.

In this section all spaces have base points and the maps are base point preserving. We shall need a result of Franks, 2.1 of [8].

Theorem. (Franks) If $g: T^r \to T^r$ is a hyperbolic toral automorphism then it is a π_1 diffeomorphism, i.e. given any homeomorphism $f: K \to K$ of a compact CW complex K and any map $h: K \to T^r$ s.t.

$$\begin{array}{ccc} \pi_1(K) & \xrightarrow{f_*} & \pi_1(K) \\ h_* \downarrow & & \downarrow h_* \\ \pi_1(T^r) & \xrightarrow[g_*]{} & \pi_1(T^r) \end{array}$$

commutes then there exists a unique map $k: K \to T^r$ homotopic to h s.t.

$$\begin{array}{ccc} K & \xrightarrow{f} & K \\ k \downarrow & & \downarrow k \\ T^r & \xrightarrow[g]{} & T^r \end{array}$$

commutes.

We start with a homeomorphism f that has a fixed point which we designate the base point. It follows from the Universal Coefficient Theorem that $H^1(M;\mathbb{Z})$ is a free abelian group $\mathbb{Z}^r$ for some r. Choose generators $\alpha_1,\ldots,\alpha_r$. We assume $f^*:H^1(M;\mathbb{Z})\to H^1(M;\mathbb{Z})$ is hyperbolic. Let G represent f^* w.r.t. the basis $\alpha_1,\ldots,\alpha_r$. G and its transpose are hyperbolic elements of $GL(r,\mathbb{Z})$. Let g be the hyperbolic toral automorphism of T^r induced by the transpose of G. Then $G=g^*:H^1(T^r;\mathbb{Z})\to H^1(T^r;\mathbb{Z})$.

Next we construct a map $h:M\to T^r$ using a suggestion of Zeeman. $H^1(M;\mathbb{Z})$ can be regarded as the homotopy classes of maps $M\to S^1$. Take a representative $h_i:M\to S^1$ of the class α_i for $i=1,\ldots,r$. Put $h=h_1\times\ldots\times h_r:M\to T^r$.

<u>Lemma 1</u>. $ghf^{-1}\simeq h$

<u>Proof</u>. Let $p_i:T^r\to S^1$ denote projection onto the i-th factor for $i=1,\ldots,r$ and use square brackets to denote homotopy classes. Then $p_ih=h_i$ so $h^*[p_i]=\alpha_i$. h^* is an isomorphism and the diagram

$$\begin{tikzcd} H^1(M;\mathbb{Z}) & H^1(M;\mathbb{Z}) \arrow[l, "f^*=G"'] \\ H^1(T^r;\mathbb{Z}) \arrow[u, "h^*"] & H^1(T^r;\mathbb{Z}) \arrow[l, "g^*=G"] \arrow[u, "h^*"'] \end{tikzcd}$$

commutes. $[p_ighf^{-1}]=f^{*-1}h^*g^*[p_i]=\alpha_i$ so $p_ighf^{-1}\simeq p_ih$ and these r homotopies can be combined to give $ghf^{-1}\simeq h$.

<u>Theorem 1</u>. Let H be the space of those base point preserving homeomorphisms of the compact manifold M whose induced map on $H^1(M;\mathbb{Z})=\mathbb{Z}^r$ is hyperbolic. Let

C be the space of continuous based maps from M to T^r homotopic to h and give H and C the C^0-topology. Then there is a continuous map $\varphi: H \to C$ s.t. the diagram

$$\begin{array}{ccc} M & \xrightarrow{f} & M \\ \varphi(f)\downarrow & & \downarrow\varphi(f) \\ T^r & \xrightarrow[g]{} & T^r \end{array}$$

commutes where $f \in H$ and g is the hyperbolic toral automorphism defined as above from $f^*: H^1(M;\mathbb{Z}) \to H^1(M;\mathbb{Z})$.

Proof. Franks' theorem and Lemma 1 guarantee the existence and uniqueness of $\varphi(f)=k$. It only remains to establish the continuity of φ at the point f. From the proof of Franks' theorem we recall that if f_1 is a homeomorphism of M close to f then

$$\varphi(f)-\varphi(f_1) = (F-I)^{-1}\{g\varphi(f)f_1^{-1}-\varphi(f)\}$$

where F is the hyperbolic automorphism $h \mapsto ghf_1^{-1}$ of the Banach space of homotopically trivial maps from M to T^r so that $F-I$ has a continuous inverse on this space. Now $\varphi(f)=g\varphi(f)f^{-1}$ so

$$\varphi(f)-\varphi(f_1) = (F-I)^{-1}\{g\varphi(f)f_1^{-1}-g\varphi(f)f^{-1}\}.$$

Since $g\varphi(f)$ is uniformly continuous we have that $\varphi(f_1)$ is close to $\varphi(f)$ when f_1 is close to f as required.

3.3 Applications to Anosov Diffeomorphisms.

If M is itself a torus then by choosing $\alpha_i=[p_i]$ we get C to be the space of maps $T^r \to T^r$ homotopic to the identity and find

Proposition 1. A hyperbolic toral automorphism g is a factor of any homeomorphism f homotopic to it.

Proof. This was realised by Franks, see Lemma (1.1) of [9], in the case where f has a fixed point. But the Lefschetz number of f is non-zero by Proposition (4.15)(a) of [29]. Now by (11.3) of [18] f has a fixed point.

Proposition 2. A hyperbolic nilmanifold automorphism g is a factor of any homeomorphism f homotopic to it.

Proof. If f has a fixed point then this proposition is an immediate corollary of Theorem (2.2) of [8] which says that g is a π_1 diffeomorphism. f always has a fixed point because its Lefschetz number is non-zero as will be proved in §4.1.

In particular if f is a diffeomorphism C^1 close to g then, by uniqueness, the quotient map taking f to g is the homeomorphism homotopic to the identity which is guaranteed by the structural stability of g. Notice how Propositions 1 and 2 overlap with the topological stability theorem of Walters [33] which says that any Anosov diffeomorphism is a factor of each homeomorphism C^0 close to it. It is natural to ask the

Question. Is it true for an arbitrary Anosov diffeomorphism g that g is a factor of any homeomorphism f homotopic to it?

If M is a nilmanifold N/D then, by proposition 1 of chapter 2, the torus T^r of theorem 1 is the maximal torus factor $N/N^1 \cdot D$. If $f_1 : N/D \to N/D$ is a hyperbolic nilmanifold automorphism then by uniqueness

$\varphi(f_1)$ must be π, the projection of the fibre bundle $\pi:N/D\to N/N^1\cdot D$. π is certainly smooth but this should not lead us to expect any differentiability in $\varphi(f_2)$ for other diffeomorphisms f_2. In fact let f_2 be C^1 close to f_1 and conjugate to it by a homeomorphism j. From the diagram

$$\begin{array}{ccc} M & \xrightarrow{f_2} & M \\ j\downarrow & & \downarrow j \\ M & \xrightarrow[f_1]{} & M \\ \pi\downarrow & & \downarrow\pi \\ T^r & \xrightarrow[g]{} & T^r \end{array}$$

we see that $\varphi(f_2)=\pi j$ which need not be differentiable. However, $\varphi(f_2)$ is still the projection of a <u>continuous</u> fibre bundle. Thus we might hope to improve theorem 1 to say that $\varphi(f)$ is a fibre bundle projection if f is Anosov but not to say that $\varphi(f)$ is differentiable.

3.4 Applications to Axiom A Diffeomorphisms.

Throughout this section f is a homeomorphism of the manifold M and $f^*:H^1(M;\mathbb{Z})\to H^1(M;\mathbb{Z})$ is assumed to be hyperbolic. It is clear that $\varphi(f)(M)$ is a connected closed g-invariant subspace of T^r. Since g is ergodic w.r.t. Haar measure either $\varphi(f)(M)$ has empty interior or it is the whole of T^r.

We consider T^r as the quotient of the r-dimensional cube I^r by the standard equivalence relation $\sim$. If $\varphi(f)$ does not map M <u>onto</u> $T^r=I^r/\sim$ then we can compose it with a retraction into $\partial I^r/\sim$ to get a map homotopic to $\varphi(f)$ and $h_1\times\ldots\times h_r$ that is into $\partial I^r/\sim$.

$T^r/(\partial I^r/\sim)$ is just S^r. Hence

Proposition 3. A sufficient condition on M for $\varphi(f)$ to be surjective is that if $h_1,\dots,h_r$ are representatives of generators of $[M,S^1]=H^1(M;\mathbb{Z})$ then the smash product $h_1\wedge\dots\wedge h_r:M\to S^r$ is not homotopically trivial.

Next we look at the situation where $\varphi(f)(M)$ is the whole of T^r.

Proposition 4. If $\varphi(f)$ is surjective so is $\varphi(f)|\Omega(f)$.

Proof. If $\varphi(f)$ is surjective and $\varphi(f)\Omega(f)$ is a proper closed subset of T^r then there is a periodic point y of g not in $\varphi(f)\Omega(f)$. Let $g^n y=y$. Then $\bigcup_{i=1}^{n}\{\varphi(f)\}^{-1}g^i y$ is a non-empty closed f-invariant set disjoint from $\Omega(f)$ which is impossible.

Proposition 5. If f is an Axiom A diffeomorphism and $\varphi(f)$ is surjective then there is a basic set Ω_1 say of f s.t. $\varphi(f)|\Omega_1$ is surjective.

Proof. Let y be a point of T^r with dense g-orbit. By Proposition 4 there is an $x\in\Omega(f)$ s.t. $\varphi(f)x=y$. Let Ω_1 be the basic set to which x belongs. Then $\varphi(f)\Omega_1$ contains the closure of the g-orbit of $\varphi(f)x$ which is T^r.

Also interesting but not new is

Proposition 6. If f satisfies Axiom A then

$$\varphi(f)W^s_f(x) \subset W^s_g(\varphi(f)x)$$

$$\text{and} \quad \varphi(f)W^u_f(x) \subset W^u_g(\varphi(f)x) \quad \text{for } x\in M.$$

3.5 Examples.

We close this chapter with some examples illustrating the results of the previous section.

Example 1. Here we take $M=T^2$ where $\varphi(f)$ is homotopic to the identity and therefore surjective. Let g be a hyperbolic automorphism of T^2 and f a DA diffeomorphism derived from g as described in [29; p.789]. (f^{-1} is considered in more detail in [35].)

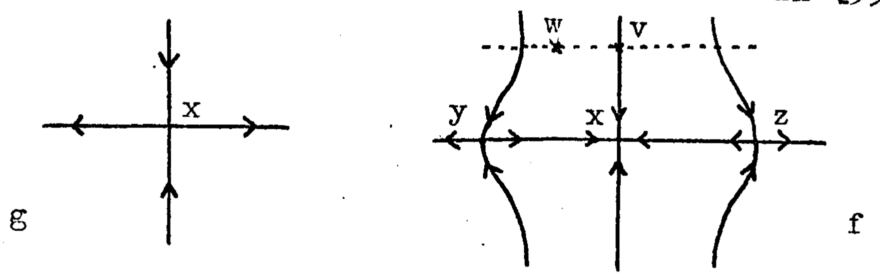

Figure 3.1

The hyperbolic fixed point 0 of g denoted by x has become three fixed points x,y,z of f. See figure 1. x is a point sink with a 2-dimensional stable manifold. The other basic set, Λ say, is 1-dimensional. The unstable manifolds are the horizontal lines in the figure just as for g except that $W_f^u(y)$ and $W_f^u(z)$ stop at x and $W_f^u(x)=\{x\}$.

By theorem 1 g is a factor of f. $\varphi(f)$ maps x,y and z to x. It maps the 2-dimensional $W_f^s(x)$ onto the vertical line $W_g^s(x)$ by a 'pinching' procedure sending a typical point w to v on the same horizontal line. Thus we cannot in general expect $\varphi(f)$ to be open. Also the 1-dimensional sets $W_f^s(y)$ and $W_f^s(z)$ are sent onto $W_g^s(x)$. But the image of $W_f^u(y)$ is only the left half of $W_g^u(x)$ so we cannot improve the $\subset$ signs to = in proposition 6 above even when the stable manifolds of f and g have the same dimension.

Λ is locally the product of a Cantor set and an interval and it is clearly Λ that satisfies proposition 5 and is mapped onto T^r. In the expanding or unstable direction where Λ is a Cantor set $W_f^u(y)\setminus\Lambda$ is a union of intervals each of which is mapped to a point. That is how the Cantor set is mapped onto an interval.

An unsolved problem in the theory of Anosov diffeomorphisms is the following. Given a hyperbolic automorphism g of T^r ($r\geq 4$) is there an Anosov diffeomorphism f homotopic to g but having $\Omega(f)\neq T^r$? If there was such an f then g would be a factor of it. Also f would have a basic set Ω_1 s.t. $\varphi(f)\Omega_1 = T^r$. Suppose the splitting of f is into j- and (r-j)-dimensional subspaces.and suppose $j\leq r-j$. By [21] $j\geq 2$. f must have a source or a sink besides Ω_1. That is f must have another basic set whose dimension is at least j. In our example 1 Λ occupies so much of T^2 that there is only really room for another basic set to be a point.

Question. Is there room in T^r for a j-dimensional basic set as well as an Ω_1 which satisfies $\varphi(f)\Omega_1 = T^r$?

Example 2. We give an example of $f:M \to M$ for which $\varphi(f)$ is not surjective on the orientable 2-dimensional manifold M of genus 2, the connected sum of two copies of T^2, T_1^2 and T_2^2 say. Let $f_1:T_1^2 \to T_1^2$ be the example 1 above with 1-dimensional source Λ_1 and point sink x_1. Let $f_2:T_2^2 \to T_2^2$ be the inverse of this diffeomorphism having a 1-dimensional sink Λ_2 and point source x_2. Now remove small discs centres x_1, x_2 and join T_1^2 and T_2^2 together at the boundaries ABCD of these discs as in figure 2. We get the manifold M and a diffeomorphism f of M that maps some points from T_1^2 into T_2^2. The basic sets of f are just the 1-dimensional source Λ_1 and the 1-dimensional sink Λ_2. This diffeomorphism was constructed in conjunction with David Chillingworth as a counterexample to a theorem of R. V. Plykin [23] that a diffeomorphism of a 2-dimensional manifold satisfying Axiom A and the no cycle property and having a 1-dimensional basic set must also have a point souce or sink.

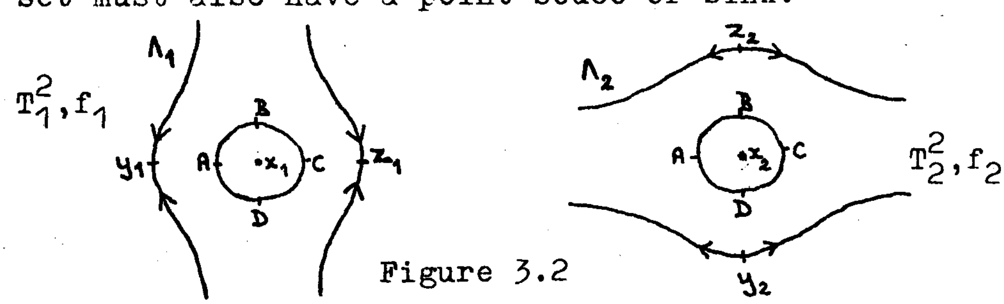

Figure 3.2

$H^1(M;\mathbb{Z}) = \mathbb{Z}^4 = (\mathbb{Z}\oplus\mathbb{Z})\oplus(\mathbb{Z}\oplus\mathbb{Z})$ and the induced map f^* is $g^*\oplus g^*$. This gives rise to the hyperbolic automorphism $g\times g$ of $T^4 = T^2 \times T^2$ and by theorem 1 $g\times g$ is a factor of f by a continuous map $\varphi(f):M \to T^4$.

We first describe a map $h:M \to T^4$ to which $\varphi(f)$

will be homotopic. h maps T_1^2 to $T^2 \times 0 \subset T^4$ and T_2^2 to $0 \times T^2 \subset T^4$ so that the circle ABCD is pinched to a point and the image hM is just the wedge of two tori, see figure 3.

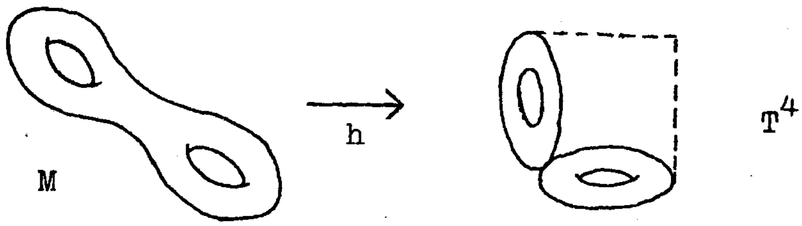

Figure 3.3

In order to understand the quotient map $\varphi(f)$ it will be helpful to see the stable and unstable manifolds in M, which are shown in figure 4 for the parts of M in figure 2. Notice how the 2-cell $W^s_{f_1}(x_1)$ in T_1^2 is foliated by stable manifolds coming into T_1^2 from T_2^2 and returning again, but that this pattern is divided in two by the exceptional lines $W^s_f(y_2)$ and $W^s_f(z_2)$ that do not return into T_2^2. Incidentally it is the need for the stable manifolds

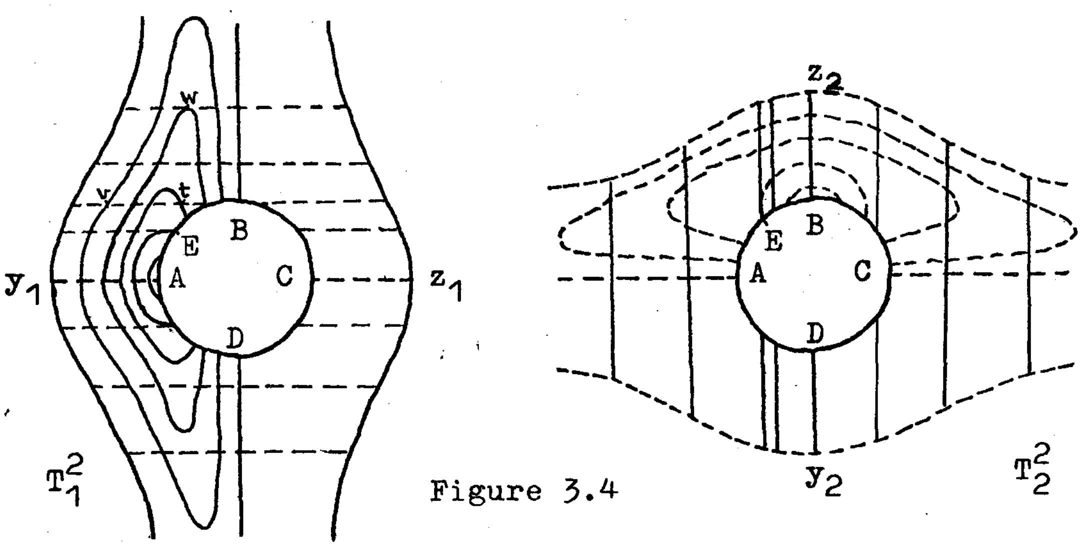

Figure 3.4

-----dashed lines represent unstable manifolds.

| solid lines represent stable manifolds and the circle ABCD.

to return into T_2^2 that ensures they have a point of non-transversal intersection, w for example, with an unstable manifold and prevents f from being an Anosov diffeomorphism.

Now we can describe $\varphi(f)$, which we shall denote by k. First recall that we have a map described in example 1 that will take Λ_1 onto $T^2\times 0$ and another taking Λ_2 onto $0\times T^2$. That deals with the nonwandering points. Notice that our map sends the fixed points y_1, z_1, y_2, z_2 to the point 0 in T^4. The wandering points of f are near Λ_1 in past time and near Λ_2 in future time. To be more precise the point t approaches the line $W_f^s(y_1)$ in past time and the line $W_f^u(z_2)$ in future time. Now $kW_f^s(y_1)$ is a line winding from 0 around the torus $T^2\times 0$ at an irrational angle. $kW_f^s(z_1)$ is the same line, and $kW_f^u(y_2)$ and $kW_f^u(z_2)$ are both a similar line winding round the torus $0\times T^2$.

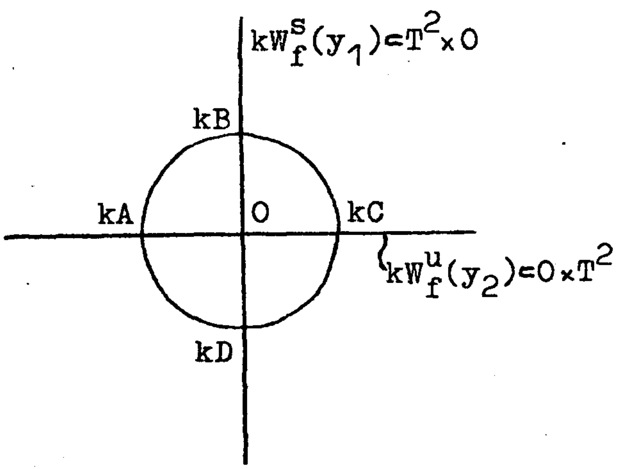

Figure 3.5

In figure 5 we draw part of the immersed plane "spanned" by $kW_f^s(y_1)$ and $kW_f^u(y_2)$. It is the projection of a plane in the universal cover R^4 of T^4 and is invariant under the automorphism $g\times g$ of T^4. The image of the circle ABCD is in this plane as shown.

Since all the wandering points of f wander from T_1^2 across to T_2^2 they must come near the circle ABCD at some time. Consider a point t of M just on the T_1^2 side of the circle ABCD. Either the stable and unstable manifolds of t both intersect the circle or this is true of ft just on the T_2^2 side of the circle. Thus kt or kft is in the plane drawn in figure 5, and by the invariance of this plane all wandering points of f are mapped into it by k. In fact the image of the wandering set is bounded by curves like hyperbolas touching k(ABCD). See figure 6. The boundary of this set is the image of points such as w or E in figure 4 where $W_f^s(w)$ and $W_f^u(w)$ meet non-transversally. The wandering points in T_1^2 are divided in two by $W_f^s(B)$ and $W_f^s(D)$ and each half is folded in two by the map k to make a crease or fin on the torus $T^2 \times 0 = k\Lambda_1$.

So the image $kM \subset T^4$ is a wedge of two tori each of which has a fin winding round it at an irrational angle and shrinking as it goes. See figure 7 where only one quarter of the fin is drawn, i.e. one quadrant of figure 6.

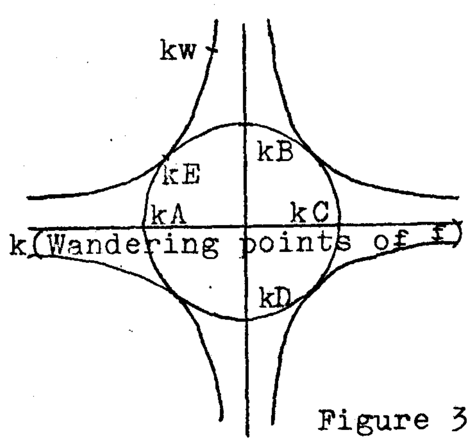

Figure 3.6

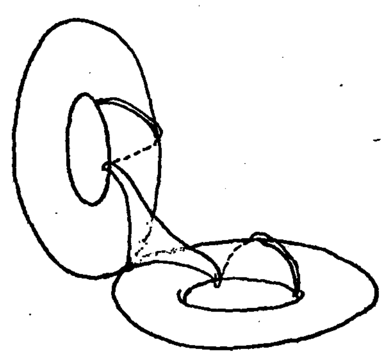

Figure 3.7

From this example we see the complexity possible in the closed g-invariant subset $\varphi(f)(M)$ of T^r. We also see that $\varphi(f)\Omega$ is not necessarily equal to $\varphi(f)(M)$ and there need not be a basic set Ω_1 with $\varphi(f)\Omega_1=\varphi(f)\Omega$. (Compare propositions 4 and 5.)

We remark that this manifold does not satisfy Hirsch's condition [14] that every infinite cyclic cover of it should have finite dimensional rational homology. So we may ask the

Question. What is the relation between Hirsch's condition and the condition of proposition 3?

Example 3. Let $g:T^2 \to T^2$ be a hyperbolic automorphism having two fixed points, 0 and t say. Modify g to f_1 as in example 1 breaking the fixed point 0 into three fixed points 0, y and z where 0 is a point sink and $y,z \in \Lambda_1$ the 1-dimensional source. For $x \in \Lambda_1$ $W^u_{f_1}(x) \cap \Omega(f_1)$ is now disconnected while $W^s_{f_1}(x) \cap \Omega(f_1)$ is connected. Now modify f_1 in a neighbourhood of the fixed point t to get a diffeomorphism f with the stable manifolds broken up too. t is now a point source of f, 0 is a point sink and there is one infinite basic set Λ say. Λ is a saddle and a Cantor set as in Smale's horseshoe example [29]. Since f is homotopic to g our theorem says that g is a factor of f and $\varphi(f)$ is homotopic to the identity and so surjective. Clearly the basic set mapped by $\varphi(f)$ onto T^2 is Λ. This shows that Ω_1 in proposition 5 need not be a source or sink.

The process of going from g to f_1 to f may be regarded as pulling first a point sink and then a point source out of $\Omega(g)=T^2$. Similarly using $\varphi(f)$ to go back from f to g may be thought of as feeding t and then O back into Λ thereby restoring to Λ its 2-dimensionality and making it the whole manifold so that g is an Anosov diffeomorphism.

Recently Smale [30] has used handle decompositions to show that any diffeomorphism is isotopic to one satisfying Axiom A with a point source, a point sink and all other basic sets 0-dimensional saddles. If we could develop a process of feeding a 0-dimensional basic set into a basic set Λ raising the dimension of Λ and get an obstruction theory for this process this would give a method of tackling the problem of which manifolds M (and which homotopy classes of diffeomorphisms of M) admit Anosov diffeomorphisms.

Chapter 4. ANOSOV DIFFEOMORPHISMS ON NILMANIFOLDS.

4.1 The Induced Map on the Fundamental Group.

In this section we calculate the Lefschetz number of an Anosov diffeomorphism of a nilmanifold $M=N/D$ and so obtain a necessary condition on the map it induces on the fundamental group D as strong as that found by Franks [9] for the torus. His result may be rephrased as

Theorem. (Franks) Let T^n be the n-dimensional torus and $f:T^n \rightarrow T^n$ an Anosov diffeomorphism. Then $f_*:\pi_1(T^n)\rightarrow\pi_1(T^n)$ has no roots of unity as eigenvalues.

We shall use the torus decomposition of M described in §2.3. M is a series of extensions of tori whose fundamental groups are D_i/D_{i-1} $i=1,\ldots,c$. Let f be a homeomorphism of M and f_* the automorphism it induces on the fundamental group D. Since we have not yet mentioned base points f_* is only defined up to an inner automorphism of D but that is sufficient for our purposes. f_* preserves the upper central series of D and so induces automorphisms $\varphi_i:D_i/D_{i-1}\rightarrow D_i/D_{i-1}$ for $i=1,\ldots,c$. (It can be shown that an inner automorphism of D induces the identity on each D_i/D_{i-1} and so the ${}_i$'s are uniquely defined.) We shall prove

Theorem 1. If f is an Anosov diffeomorphism then none of the φ_i's have a root of unity as an eigenvalue.

As we remarked in §3.1 Hirsch [14] proved this for the map induced by f on $H_1(M;R)$. Our proof uses a spectral sequence to calculate the Lefschetz number of f and shows the remarkable fact that it is independent of the twists with which the tori are put together to make up M.

Proof. Choose an automorphism of D induced by f. By fact 2 of §2.1 it extends uniquely to an automorphism $G:N \to N$ which induces a nilmanifold automorphism g of N/D. The diffeomorphisms f,g induce conjugate automorphisms of the fundamental group D and so by fact 3 of §2.1 induce the same map of $H_*(M)$. Therefore $L(f)=L(g)$.

The automorphism φ_i of the fundamental group of the i-th torus of M is induced by an automorphism g_i say of this torus and $g:M \to M$ is the extension of g_1 by g_2 by ... by g_c. We show that $L(g)=L(g_1 \times \ldots \times g_c)$. A special case of this was noticed by Bowen [3; p.395]. In fact it follows from the next lemma by induction on c and the observation that the condition about trivial action is satisfied because the series $\{D_i\}$ is central.

Lemma 1. Let $\pi:X,* \to B,*$ be a fibre bundle with fibre $F=\pi^{-1}*$ and suppose $\pi_1(B)$ acts trivially on the homology of F. Assume that at least one of B,F is compact. Let (ψ,χ) be a bundle map i.e. a pair of continuous maps s.t. the diagram

$$\begin{array}{ccc} X,* & \xrightarrow{\psi} & X,* \\ \pi\downarrow & & \downarrow\pi \\ B,* & \xrightarrow[\chi]{} & B,* \end{array}$$

commutes and let $\omega=\psi|F$. Then $L(\psi)=L(\chi\times\omega)$.

<u>Remark</u>. $\psi:X\to X$ and $\chi\times\omega:B\times F\to B\times F$ differ by twists in the fibres so the lemma says that the Lefschetz number ignores these twists. If $\psi=id_X$ then the result reduces to the multiplicative property of the Euler characteristic (theorem 9.3.1 of [32]) which, however, is true without the condition of trivial action. This condition is required here since the Klein bottle K is an S^1 bundle over S^1 failing to satisfy it and the map $\psi:K\to K$ that induces the identity in the fibre but wraps the base three times round itself has Lefschetz number -2 but the corresponding map of T^2 has Lefschetz number 0.

<u>Proof of Lemma 1</u>. We use cubical singular homology with real coefficients and the Serre spectral sequence, see [24] and [13]. Let ${}^{0}_{n}\square(X)$ be the real vector space with basis all maps of the standard n-cube I^n into X such that all vertices are mapped to $*$. Filter ${}^{0}_{n}\square(X)$ as follows. Take a basis element $\sigma\epsilon{}^{0}_{n}\square(X)$, $\sigma:I^n\to X$ and define p to be the least integer such that $\pi\sigma(u_1,\dots,u_n)$ is independent of $u_{p+1},\dots,u_n$. Then $\sigma\epsilon{}^{0}_{n}\square_p(X)$. Now $\psi:X,*\to X,*$ induces a chain map of ${}^{0}_{n}\square(X)$ to itself which preserves the filtration by p. So ψ induces a map which we denote by ψ_* on every term E^r_{pq} of the spectral sequence obtained from ${}^{0}_{n}\square_p$.

Define

$$L(\psi,E^r) = \sum_{p,q}(-1)^{p+q}\mathrm{trace}(\psi_*:E^r_{pq}\to E^r_{pq}).$$

The Hopf Trace Theorem, see e.g. 5.1.18 of [13], says that the Lefschetz number of a chain map of a finitely generated chain group is the same as the Lefschetz number of the induced map on its homology groups. E^{r+1} is defined as the homology of the chain group E^r so $L(\psi,E^r) = L(\psi,E^{r+1})$.

Now $E^2_{pq} = H_p(B;H_q(F)) = H_p(B)\otimes H_q(F)$ by the assumption of trivial action and $H_n(B\times F) = \bigoplus_{p+q=n}H_p(B)\otimes H_q(F)$ by the Kunneth formula. So $L(\psi,E^2) = L(\chi\times\omega)$.

Since one of B,F is compact there is an m such that $E^m_{pq} = E^\infty_{pq}$. Then

$$\begin{aligned} L(\psi,E^\infty) &= \textstyle\sum_{p,q}(-1)^{p+q}\operatorname{trace}(\psi_*:E^\infty_{pq}\to E^\infty_{pq}) \\ &= \textstyle\sum_n(-1)^n\operatorname{trace}(\psi_*:H_n(X)\to H_n(X)) \\ &= L(\psi). \end{aligned}$$

Therefore

$$L(\psi) = L(\psi,E^\infty) = L(\psi,E^m) = L(\psi,E^2) = L(\chi\times\omega).$$

<u>Completion of Proof of Theorem 1</u>. Now we can calculate

$$L(f) = L(g) = L(g_1\times\ldots\times g_c) = \prod(1-\lambda)$$

where the product is taken over all eigenvalues λ counted with multiplicity of all the maps φ_i [29; p.769]. If one of these eigenvalues is a jth root of unity then $L(f^{jr})=0$ for any r according to this calculation. But some f^{jr} must have a fixed point. So $L(f^{jr})\neq 0$ if we can show that all the fixed points of f^{jr} have the same Lefschetz index. This is easy if the expanding bundle E^u is orientable, see [9; p.123]. Moreover if E^u is not orientable we can use the same trick as Franks. Namely we construct a covering f' of f^{jr} on the covering space of M corresponding to that subgroup

H of $D=\pi_1(M)$ which is the inverse image of $2D/[D,D]$ under the Hurewicz map $D \to D/[D,D]=H_1(M;\mathbf{Z})$. Then f' is an Anosov diffeomorphism with orientable expanding bundle so the map induced by f' on H and hence the map induced by f on D has no eigenvalues which are roots of unity. This completes the proof of theorem 1.

<u>Corollary 1</u>. A hyperbolic nilmanifold automorphism g has $L(g) \neq 0$.

<u>Proof</u>. $L(g)=\prod(1-\lambda)$ so $L(g)=0$ implies some $\lambda=1$. Alternatively, g has the fixed point eD and its expanding bundle is orientable so $L(g)\neq 0$. This corollary was needed for proposition 2 of chapter 3.

4.2 <u>Summary of What is Known About These Diffeomorphisms</u>.

In Franks' investigation of Anosov diffeomorphisms on tori [9] he also proved the

<u>Theorem</u>. (Franks) If $f:T^n \to T^n$ is an Anosov diffeomorphism with $\Omega(f)=T^n$ and if $f_*:H_1(T^n;R) \to H_1(T^n;R)$ is hyperbolic then f is topologically conjugate to a hyperbolic toral automorphism.

<u>Theorem 2</u>. If $f:N/D \to N/D$ is an Anosov diffeomorphism with $\Omega(f)=N/D$ and inducing a hyperbolic automorphism (i.e. one for which the φ_i's have no eigenvalues of modulus one) on the fundamental group D then f is topologically conjugate to a hyperbolic nilmanifold automorphism.

<u>Proof</u>. Define g to be a hyperbolic nilmanifold automorphism homotopic to f as in the proof of

theorem 1. Proposition 2 of chapter 3 says that g is a factor of f by a continuous map, k say, homotopic to the identity. Now (1.5) to (1.8) of [9] go through as in the torus case to prove that k is a local homeomorphism and hence a homeomorphism.

Putting together theorems 1 and 2 of this chapter we see that the open questions about Anosov diffeomorphisms of nilmanifolds are the same as those for tori:

(1) Is there an Anosov diffeomorphism f of N/D whose induced map on the fundamental group is hyperbolic but with nonwandering set not the whole manifold?

(2) Is there an Anosov diffeomorphism f of N/D whose induced map on the fundamental group has an eigenvalue of modulus one but not a root of unity?

4.3 Hyperbolic Automorphisms of Nilmanifolds.

If the two open questions of the previous section could be answered in the negative the only work remaining in the classification of Anosov diffeomorphisms of nilmanifolds would be to find their hyperbolic automorphisms. By facts 1 and 2 of §2.1 that means find the hyperbolic automorphisms of finitely generated torsion-free nilpotent groups. Now an automorphism φ of D breaks down as in §4.1 into c automorphisms φ_i of $D_i/D_{i-1} = \mathbb{Z}^{r_i}$ for $i=1,\dots,c$. So we can rephrase our question as follows. Given $r_1,\dots,r_c$ and hyperbolic elements $\varphi_i \in GL(r_i,\mathbb{Z})$ for $i=1,\dots,c$ which of the possible extensions of $\mathbb{Z}^{r_1}$ by

$\mathbb{Z}^{r_2}$ by ... by $\mathbb{Z}^{r_c}$ admit an automorphism built from $\varphi_1,\ldots,\varphi_c$?

§4.1 says that as far as the number of periodic points is concerned it does not matter with what twists the tori are put together to make the nilmanifold. Here we ask how much information is lost by this approach, i.e. what twists were possible for particular automorphisms of the tori. (In [2a] Auslander and Scheuneman investigated hyperbolic automorphisms of nilmanifolds N/D but purely in terms of automorphisms of the Lie algebra of N fixing a "$\mathbb{Z}$-subalgebra" of it.)

An <u>extension</u> of the group A by the group B is defined to be an exact sequence

$$1 \to A \xrightarrow{i} G \xrightarrow{j} B \to 1$$

where 1 denotes the group with only one element. Since D was broken down by its upper central series we shall only be concerned with <u>central</u> extensions, that is where iA is in the centre of G. In particular A must be abelian. The central extensions of A by B are in one-one correspondence with the elements of $H^2(B;A)$, see [17; p.212] for example.

Let B be any group and A an abelian group. Let α,β be automorphisms of A and B respectively. Which extensions D of A by B admit an automorphism δ s.t. the following diagram commutes?

$$\begin{array}{ccccccccc} 1 & \to & A & \to & D & \to & B & \to & 1 \\ \downarrow & & \downarrow{\scriptstyle\alpha} & & \downarrow{\scriptstyle\delta} & & \downarrow{\scriptstyle\beta} & & \downarrow \\ 1 & \to & A & \to & D & \to & B & \to & 1 \end{array} \qquad (1)$$

(We shall say that δ is an extension of α by β.)

By [17; p.214] these extensions are precisely the ones corresponding to elements $d \in H^2(B;A)$ for which $\alpha_* d = \beta^* d$ where α_*, β^* are the automorphisms of $H^2(B;A)$ induced by $\alpha: A \to A$, $\beta: B \to B$. If d and d' satisfy this condition so does $md+m'd'$ for $m,m' \in \mathbb{Z}$. Thus we get

<u>Proposition 1</u>. The central extensions of A by B admitting an automorphism δ making the diagram (1) commute form a sub $\mathbb{Z}$-module of $H^2(B;A)$.

To use this condition for building groups D of nilpotency class (the length of the upper central series) c and hyperbolic automorphisms of them involves calculating H^2 of groups like D which, even with the technique of the spectral sequence of a group extension, is very heavy going. But the condition is certainly useful for groups of nilpotency class 2.

For such groups D we consider central extensions of $\mathbb{Z}^a$ by $\mathbb{Z}^b$ for positive integers a,b. What is $H^2(\mathbb{Z}^b;\mathbb{Z}^a)$? First consider $H^2(\mathbb{Z}^b;\mathbb{Z})$. The cohomology of a group G is isomorphic to the cohomology of an Eilenberg-MacLane space K(G,1). This is either taken as the definition of $H^*(G)$ or deduced from the abstract definition as in [13; p.461]. The torus T^b is a $K(\mathbb{Z}^b,1)$ so $H^2(\mathbb{Z}^b;\mathbb{Z}) = (\mathbb{Z}^b \wedge \mathbb{Z}^b) \pitchfork \mathbb{Z}$ where $\wedge$ means the exterior product and $G \pitchfork H$ means the group of homomorphisms from G to H. It follows that $H^2(\mathbb{Z}^b;\mathbb{Z}^a) = (\mathbb{Z}^b \wedge \mathbb{Z}^b) \pitchfork \mathbb{Z}^a$. This can be regarded as the group of skew-symmetric homomorphisms $\mathbb{Z}^b \times \mathbb{Z}^b \to \mathbb{Z}^a$.

Now take hyperbolic elements $\alpha\in GL(a,\mathbb{Z})$ and $\beta\in GL(b,\mathbb{Z})$. Let $\lambda_1,\dots,\lambda_a$ be the eigenvalues of α and $\mu_1,\dots,\mu_b$ those of β (counting multiplicity). The eigenvalues of $\beta^{*-1}\alpha_*:(\mathbb{Z}^b\wedge\mathbb{Z}^b)\pitchfork\mathbb{Z}^a\to(\mathbb{Z}^b\wedge\mathbb{Z}^b)\pitchfork\mathbb{Z}^a$ are $\lambda_i\mu_j^{-1}\mu_k^{-1}$, $1\leqslant i\leqslant a$, $1\leqslant j<k\leqslant b$. For example, if α and β are both diagonalizable and $x_1,\dots,x_a;y_1,\dots,y_b$ are bases of eigenvectors in $\mathbb{C}^a,\mathbb{C}^b$ then the eigenvalue $\lambda_i\mu_j^{-1}\mu_k^{-1}$ of $\beta^{*-1}\alpha_*$ corresponds to the eigenvector $d\in(\mathbb{C}^b\wedge\mathbb{C}^b)\pitchfork\mathbb{C}^a$ defined by $d(y_j,y_k)=x_i$, $d(y_k,y_j)=-x_i$, d sends all other pairs of basis vectors to zero. This is because $(\beta^{*-1}\alpha_*(d))(y_j,y_k) = \alpha d(\beta^{-1}y_j,\beta^{-1}y_k)$ $= \mu_j^{-1}\mu_k^{-1}\alpha d(y_j,y_k) = \mu_j^{-1}\mu_k^{-1}\alpha(x_i) = \lambda_i\mu_j^{-1}\mu_k^{-1}d(y_j,y_k)$ so that $\beta^{*-1}\alpha_* d = \lambda_i\mu_j^{-1}\mu_k^{-1}d$.

We are interested in fixed points of $\beta^{*-1}\alpha_*$ so we want this transformation to have an eigenvalue 1. If $\lambda_i\mu_j^{-1}\mu_k^{-1} = 1$ for some i,j,k, $j\neq k$, then a corresponding eigenvector will have rational coordinates so some multiple of it will have integer coordinates. Thus we get

<u>Proposition 2.</u> If $\alpha\in GL(a,\mathbb{Z})$ and $\beta\in GL(b,\mathbb{Z})$ are hyperbolic matrices, $p(\alpha)$ and $p(\beta)$ are their characteristic equations and there are two roots of $p(\beta)$ whose product is a root of $p(\alpha)$ then there is a non-toral $(a+b)$-dimensional two-step nilmanifold supporting an Anosov diffeomorphism which induces on the fundamental group an extension of α by β. All hyperbolic automorphisms of two-step nilmanifolds are obtained in this way by varying α,β,a and b. If on the other hand there are no two roots of $p(\beta)$ whose

product is a root of $p(\alpha)$ only the trivial extension is possible, that is only the hyperbolic automorphism induced by $\alpha\times\beta$ on the torus T^{a+b}.

Is the non-toral nilmanifold in Proposition 2 unique or are there many such? To answer this question we shall need a lemma.

Lemma 2. Let H be a subgroup of the free abelian group $A=\mathbf{Z}^a$ so that rH is also a subgroup of A for any $r\in\mathbf{Z}$. Then among the quotient groups $A/(rH)$, $r\in\mathbf{Z}$, there are infinitely many non-isomorphic groups.

Proof. It suffices to show that, having constructed $s_q>s_{q-1}>\dots>s_1>0$ so that all the groups $A/(s_iH)$ are non-isomorphic we can construct $s_{q+1}>s_q$. The result will then follow by induction on q since the induction can be started with $s_1=1$. Choose $x\in A$ s.t. $x+s_qH$ is an element of largest possible finite order, m say, in $A/(s_qH)$. $mx\in s_qH$ but if $0<j<m$ then $jx\notin s_qH$.

Case 1. $m=1$. $x\in s_qH$ and $A/(s_qH)$ has trivial torsion subgroup. Choose $y\in s_qH$ s.t. $\nexists z\in s_qH$ with $y=2z$. Then $y\notin 2s_qH$ and so $A/(2s_qH)$ has non-trivial torsion subgroup. Put $s_{q+1}=2s_q$.

Case 2. $m>1$. Put $s_{q+1}=ms_q$. $x\notin s_qH$ so $mx\notin s_{q+1}H$. Also if $0<j<m$ $jx\notin s_{q+1}H\subset s_qH$. Therefore $x+s_{q+1}H$ has order $>m$ but $\leq m^2$. Since the maximal order of the torsion elements of $A/(s_qH)$ increases monotonically with q all these groups are non-isomorphic.

This lemma is needed for

Proposition 3. If N/D is an (a+b)-dimensional two-step nilmanifold with $\mathbb{Z}^a$ as the centre of D and N/D admits a hyperbolic nilmanifold automorphism given by an automorphism of D that is an extension of α by β for some fixed hyperbolic elements $\alpha\in GL(a,\mathbb{Z})$, $\beta\in GL(b,\mathbb{Z})$ then there is a countably infinite set of non-homeomorphic nilmanifolds with the same properties.

Proof. There is an extension $1\to\mathbb{Z}^a\xrightarrow{i}D\xrightarrow{j}\mathbb{Z}^b\to 1$, where $i\mathbb{Z}^a=D_1$ the centre of D, given by a $\beta^{*-1}\alpha_*$-invariant element $d\in H^2(\mathbb{Z}^b;\mathbb{Z}^a)$. For any $r\in\mathbb{Z}$ rd is also a $\beta^{*-1}\alpha_*$-invariant element of $H^2(\mathbb{Z}^b;\mathbb{Z}^a)$. So rd corresponds to a central extension D(r), say, of $\mathbb{Z}^a$ by $\mathbb{Z}^b$ that admits an automorphism which is an extension of α by β. It will be sufficient to find an infinite number of non-isomorphic groups among these D(r).

As before let $D(r)_1$ denote the centre of D(r) and $D(r)^1$ that subgroup of D(r) generated by all elements of the form $p^{-1}q^{-1}pq$ for $p,q\in D(r)$. The inclusions between the <u>upper</u> and <u>lower</u> central series of D(r) can now be displayed as

$$\begin{array}{ccccc} \{e\} & \subset & D(r)_1 & \subset & D(r) \\ \| & & \cup & & \| \\ \{e\} & \subset & D(r)^1 & \subset & D(r) \end{array} .$$

We shall investigate $D(r)_1/D(r)^1$. (This same quotient group was considered in [2] for 3-dimensional nilmanifolds.)

To calculate the groups $D(r)_1$, $D(r)^1$ we must first explain how the group D(r) is defined from the element rd of $H^2(\mathbb{Z}^b;\mathbb{Z}^a)$. The underlying set of D(r) is

$\{(x,u); x\in\mathbb{Z}^a,\ u\in\mathbb{Z}^b\}$. The group operation on this set is defined using a cocycle representing the cohomology class rd. A 2-cochain is a function (not necessarily a homomorphism) from $\mathbb{Z}^b\times\mathbb{Z}^b\to\mathbb{Z}^a$. Such a cochain is a cocycle precisely when the product we now define is associative. Let $\sigma:\mathbb{Z}^b\times\mathbb{Z}^b\to\mathbb{Z}^a$ be a cocycle representing d and use the cocycle $r\sigma$ to represent rd. The product in D(r) of elements (x,u) and (y,v) is defined to be

$$(x+y+r\sigma(u,v),u+v).$$

This is equal to $(y,v)(x,u) = (y+x+r\sigma(v,u),v+u)$ if and only if $r\sigma(u,v)=r\sigma(v,u)$. But this is equivalent to $\sigma(u,v)=\sigma(v,u)$. Thus (x,u) commutes with all elements in D(r) if and only if it does in D. $\therefore D(r)_1\cong D_1=\mathbb{Z}^a$.

Now $(x,u)^{-1} = (-x-r\sigma(u,-u),-u)$. So

$$(x,u)^{-1}(y,v)^{-1}(x,u)(y,v) =$$

$$(-r\sigma(u,-u)-r\sigma(v,-v)+r\sigma(-u,-v)+r\sigma(u,v)+r\sigma(-u-v,u+v),0)$$

But this generator of $D(r)^1$ is just r times a generator of D^1. Thus $D(r)^1=rD^1$. Now, by Lemma 2, there are infinitely many groups D(r) with non-isomorphic $D(r)_1/D(r)^1$. Hence there are infinitely many non-isomorphic groups D(r) and so the corresponding nilmanifolds are non-homeomorphic.

A question not investigated here is what automorphisms δ are possible in the commutative diagram

$$\begin{array}{ccccccccc}
1 & \to & \mathbb{Z}^a & \to & D & \to & \mathbb{Z}^b & \to & 1 \\
\downarrow & & \downarrow{\scriptstyle\alpha} & & \downarrow{\scriptstyle\delta} & & \downarrow{\scriptstyle\beta} & & \downarrow \\
1 & \to & \mathbb{Z}^a & \to & D & \to & \mathbb{Z}^b & \to & 1
\end{array}$$

for fixed D, α, β. It would seem that there are many such automorphisms not necessarily conjugate to each other in the group Aut(D) and [17; p.216] connects this question with the group $H^1(\mathbb{Z}^b;\mathbb{Z}^a) = \mathbb{Z}^b \wedge \mathbb{Z}^a$.

Smale's Example. We illustrate this section now by showing how proposition 2 applies to the two hyperbolic automorphisms Smale defined on a certain 6-dimensional two-step nilmanifold in [29; p.762]. In this example he defines a 6-dimensional nilpotent Lie group G and a uniform discrete subgroup Γ of it. Γ is the group we have called D. Smale works mainly with the subset $\Gamma_0 = \exp^{-1}\Gamma$ of the Lie algebra of G. This is a set of matrices with entries in the field $Q(\sqrt{3})$. σ denotes the automorphism of $Q(\sqrt{3})$ sending $\sqrt{3}$ to $-\sqrt{3}$. Γ_0 is the set of 6×6 matrices $\begin{pmatrix} P & 0 \\ 0 & P^\sigma \end{pmatrix}$ where $P = \begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix}$; x,y,z have the form $m+n\sqrt{3}$ for $m,n\in\mathbb{Z}$ and P^σ denotes the result of applying σ to each entry of P.

An element of Γ_0 is clearly determined by the matrix P. We define generators for Γ by specifying the matrix P for the corresponding elements of Γ_0.

J_1, J_2 have $x=1$ and $\sqrt{3}$ respectively and other entries 0.
K_1, K_2 " $y=1$ " $\sqrt{3}$ " " " " ".
L_1, L_2 " $z=1$ " $\sqrt{3}$ " " " " ":

Now only L_1 and L_2 have bracket zero with all these generators so the centre of Γ (which we call Γ_1) is generated by $\{\exp L_1, \exp L_2\}$. Γ/Γ_1 is generated by the cosets of the remaining generators of Γ, namely

$\{\exp J_1, \exp J_2, \exp K_1, \exp K_2\}$. So this nilmanifold is a principal fibre bundle with fibre T^2 and base T^4.

Now let S be the hyperbolic matrix $\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \in GL(2,\mathbb{Z})$ whose eigenvalues are $\lambda = 2+\sqrt{3}$ and $\lambda^{-1} = 2-\sqrt{3}$. Smale's first hyperbolic automorphism induces $\alpha_1 = S^3$ on Γ_1 and $\beta_1 = \begin{pmatrix} S & 0 \\ 0 & S^2 \end{pmatrix}$ on Γ/Γ_1 both matrices w.r.t. the generators described above. β_1 has eigenvalues $\lambda, \lambda^{-1}, \lambda^2, \lambda^{-2}$ and α_1 has λ^3, λ^{-3} so the conditions of proposition 2 are clearly satified. Smale's second hyperbolic automorphism induces $\alpha_2 = S^{-2}, \beta_2 = \begin{pmatrix} S & 0 \\ 0 & S^{-3} \end{pmatrix}$ so we have two eigenvalues of β_2 whose product is an eigenvalue of α_2.

4.4 Zeta Functions.

The reference for this section is §I.4 of [29]. The false zeta function of a diffeomorphism f, $\tilde{\zeta}(f,t)$, is defined by

$$\tilde{\zeta}(f,t) = \exp \sum_{m=1}^{\infty} (1/m) L(f^m) t^m .$$

If the Lefschetz index of all the fixed points of f^m is the same then $N_m(f) = |L(f^m)|$ and $\zeta(f)$ is easily calculated from $\tilde{\zeta}(f)$. In §4.1 we found that an Anosov diffeomorphism f of a nilmanifold has the same Lefschetz number as a certain associated automorphism of the torus of the same dimension. Thus if $\lambda_1, \ldots, \lambda_n$ are all the eigenvalues of all the φ_i's counted with multiplicity then

$$\tilde{\zeta}(f,t) = \prod (1 - \lambda_{i_1} \lambda_{i_2} \cdots \lambda_{i_k} t)^{(-1)^{k+1}}$$

where the product is taken over all $(i_1, \ldots, i_k)$ s.t.

$1 \leq i_1 < i_2 < \ldots < i_k \leq n$.

Suppose now that g is a hyperbolic nilmanifold automorphism so that $|\lambda_i| \neq 1$ for each i. $\zeta(g)$ depends on $\tilde{\zeta}(g)$, the dimension of the expanding bundle E^u and whether Dg preserves or reverses the orientation of E^u. These last two can be calculated from those λ_i with $|\lambda_i| > 1$ and are the same for $g_1 \times \ldots \times g_c$. These remarks prove

Proposition 4. $\zeta(g) = \zeta(g_1 \times \ldots \times g_c)$.

Thus the zeta function does not distinguish between the nilmanifold automorphism and the associated toral automorphism. But this does not detract from its power to distinguish between diffeomorphisms of the same manifold.

$g: N/D \longrightarrow N/D$ is covered by an automorphism $G: N \longrightarrow N$ which induces an automorphism of the Lie algebra of N. The eigenvalues of this automorphism are just $\lambda_1, \ldots, \lambda_n$. If the Lie algebra is not abelian then there must be eigenspaces corresponding to eigenvalues λ_i, λ_j say whose bracket is not zero making $\lambda_i \lambda_j$ an eigenvalue too. Hence

Proposition 5. The zeta function above of a product of toral automorphisms can only be the zeta function of a non-toral nilmanifold automorphism if a factor for which $k=1$ cancels with a factor for which $k=2$.

Question. Can more information about which factors of $\zeta(g_1 \times \ldots \times g_c)$ cancel determine which nilmanifolds admit an automorphism corresponding to $g_1 \times \ldots \times g_c$?

REFERENCES.

1. D. V. Anosov, "Geodesic flows on closed Riemann manifolds with negative curvature", Proc. Steklov Inst. Math., 90 (1967). MR 36#7157.

2. L. Auslander, L. Green and F. Hahn, "Flows on homogeneous spaces", Annals of Maths Studies, No. 53, Princeton University Press, Princeton N. J., 1963. MR 29#4841.

2a. L. Auslander and J. Scheuneman, "On certain automorphisms of nilpotent Lie groups", Global Analysis, Proc. Symp. Pure Math., 14 (1970), 9-15. MR 42#5284.

3. R. Bowen, "Periodic points and measures for axiom A diffeomorphisms", Trans. Amer. Math. Soc., 154 (1971), 377-397.

4. R. Bowen, "Markov partitions for axiom A diffeomorphisms", Amer. J. Math., 92 (1970), 725-747. MR 43#2740.

5. R. Bowen, "Markov partitions and minimal sets for axiom A diffeomorphisms", Amer. J. Math., 92 (1970), 907-918. MR 43#2739.

6. R. Bowen, "Symbolic dynamics for hyperbolic flows", Preprint 1972.

7. R. Bowen and O. E. Lanford III, "Zeta functions of restrictions of the shift transformation", Global Analysis, Proc. Symp. Pure Math., 14 (1970), 43-49. MR 42#6284.

8. J. Franks, "Anosov diffeomorphisms", Global Analysis, Proc. Symp. Pure Math., 14 (1970), 61-93. MR 42#6871.

9. J. Franks, "Anosov diffeomorphisms on tori", Trans. Amer. Math. Soc., 145 (1969), 117-124. MR 40#6567.

10. J. Franks, "Differentiably Ω-stable diffeomorphisms", Topology, 11 (1972), 107-114.

11. J. Guckenheimer, "Axiom A + no cycles $\Rightarrow \zeta_f(t)$ rational", Bull. Amer. Math. Soc., 76 (1970), 592-594. MR 41#8078.

12. J. Guckenheimer, "Absolutely Ω-stable diffeomorphisms", Topology, 11 (1972), 195-199.

13. P. Hilton and S. Wylie, "Homology theory: an introduction to algebraic topology", Cambridge University Press, Cambridge, 1960. MR 22#5963.

14. M. Hirsch, "Anosov maps, polycyclic groups and homology", Topology, 10 (1971), 177-183.

15. M. Hirsch and C. Pugh, "Stable manifolds and hyperbolic sets", Global Analysis, Proc. Symp. Pure Math., 14 (1970), 133-163. MR 42 6872.

16. N. Jacobson, "Lie algebras", Interscience, New York, 1962. MR 26#1345.

17. S. Lang, "Rapport sur la cohomologie des groupes", Benjamin, New York, 1967. MR 35#2948.

18. S. Lefschetz, "Introduction to topology", Princeton University Press, Princeton N. J., 1949. MR 11,193.

19. A. I. Malcev, "On a class of homogeneous spaces", Amer. Math. Soc. Transl., (1) 9 (1962), 276-307. MR 10,507.

20. A. I. Malcev, "Nilpotent torsion-free groups" Russian, Izv. Akad. Nauk SSSR Ser. Mat., 13 (1949), 201-212. MR 10,507.

21. S. Newhouse, "On codimension one Anosov diffeomorphisms", Amer. J. Math., 92 (1970), 761-770. MR 43#2741.

22. W. Parry, "Ergodic properties of affine transformations and flows on nilmanifolds", Amer. J. Math., 91 (1969), 757-771. MR 41#5595.

23. R. V. Plykin, "The topology of basis sets for Smale diffeomorphisms", Math. USSR Sbornik, 13 (1971), 297-307. Translated from Mat. Sbornik, 84 (126) (1971), 301-312.

24. J.-P. Serre, "Homologie singulière des éspaces fibrés", Ann. Math., (2) 54 (1951), 425-505. MR 13,574.

25. M. Shub, "Endomorphisms of compact differentiable manifolds", Amer. J. Math., 91 (1969), 175-199. MR 39#2169.

27. C. P. Simon, "On a classification of a Baire set of diffeomorphisms", Bull. Amer. Math. Soc., 77 (1971), 783-787.

28. Ya. G. Sinai, "Markov partitions and C-diffeomorphisms", Func. Anal. and its Appl., 2 (1968), 64-89. MR 38#1361.

29. S. Smale, "Differentiable dynamical systems", Bull. Amer. Math. Soc., 73 (1967), 747-817. MR 37#3598.

30. S. Smale, "Stability and isotopy in discrete dynamical systems", Dynamical Systems Conference, Rio de Janeiro, 1971.

31. S. Smale and R. Abraham, "Non-genericity of Ω-stability", Global Analysis, Proc. Symp. Pure Math., 14 (1970), 5-8. MR 42#6867.

32. E. Spanier, "Algebraic topology", McGraw-Hill, New York, 1966. MR 35#1007.

33. P. Walters, "Anosov diffeomorphisms are topologically stable", Topology, 9 (1970), 71-78. MR 40#8069.

34. R. F. Williams, "The zeta function in global analysis", Global Analysis, Proc. Symp. Pure Math., 14 (1970), 335-339. MR 42#1159.

35. R. F. Williams, "The DA maps of Smale and structural stability", Global Analysis, Proc. Symp. Pure Math., 14 (1970), 329-334. MR 41#9296.