Flow equivalence, hyperbolic systems and a new zeta function for flows.

Comment. Math. Helvetici 57 (1982) 237–259 0010–2571/82/002237-23$01.50+0.20/0

Flow equivalence, hyperbolic systems and a new zeta function for flows

DAVID FRIED*

Abstract. We analyze the dynamics of diffeomorphisms in terms of their suspension flows. For many Axiom A diffeomorphisms we find simplest representatives in their flow equivalence class and so reduce flow equivalence to conjugacy. The zeta functions of maps in a flow equivalence class are correlated with a zeta function ζ_H for their suspended flow. This zeta function is defined for any flow with only finitely many closed orbits in each homology class, and is proven rational for Axiom A flows. The flow equivalence of Anosov diffeomorphisms is used to relate the spectrum of the induced map on first homology to the existence of fixed points. For Morse–Smale maps, we extend a result of Asimov on the geometric index.

0. Introduction

Since Poincare's time, the dynamical properties of a smooth differential equation have been studied in terms of maps between local transversals to the flow ρ. Near a periodic orbit γ, for instance, one may introduce a small transverse disc D of codimension one which cuts γ once and study the partially defined map $f: D \to D$ obtained by following the flowline through $d \in D$ until it again meets D in $f(d) = \rho_t d$, $t > 0$. This local section allows one to study the stability of γ (or other properties of ρ near γ) in terms of the stability (etc.) of the fixed point $\gamma \cap D$ for f.

To obtain a global picture of a smooth (C^∞, say) non-singular flow ρ on a compact manifold M one might cover M with flowboxes (i.e., a disjoint family of transverse, codimension-one, discs D_i where $B_i = \{\rho_t D_i : 0 \le t \le \eta_i\}$ is an embedded cylinder and $\{B_i\}$ covers M) and analyze the measurable first-return map f, defined as above. While this reduction is satisfactory for the purposes of ergodic theory, the discontinuity of f makes this approach useless for studying the topological behavior of ρ.

For a satisfactory global reduction of ρ to a mapping, one needs a *cross-section*, that is a transverse, closed submanifold K which meets every flowline.

*Partially supported by MCS 76-08795.

When such a K exists, the first return map $f(k)=\rho_{t(k)}(k)$ (where $t(k)>0$, $\rho_t(k)\notin K$ for $0<t<t(k)$ and $\rho_{t(k)}(k)\in K$) is a well-defined diffeomorphism and the recurrence properties of f accurately reflect those of ρ. We call a flow *circular* if it admits a cross-section.

Not all nonsingular flows are circular (necessary and sufficient conditions are given in section 1) so that this pleasant reduction is often impossible. However any smooth map $t:K\to(0,\infty)$ and diffeomorphism f of K determine the *suspended flow* ρ with cross-section K, first-return map f and time of first return t. On the manifold $M=\{(k,s)\mid 0\le s\le t(k)\}/(k,t(k))=(f(k),0)$ the flow ρ is induced by d/ds.

Even when cross-sections exist, there may well be many essentially distinct cross-sections and first-return maps. One calls the resulting relation amongst first-return maps *flow equivalence*, so f and g are flow equivalent if they have conjugate suspended flows. For example, all the rational flows on the torus are conjugate and so all rational rotations of S^1 are flow equivalent. This means that it is difficult to decide whether two flows with cross-section are conjugate: it does not suffice generally to investigate the conjugacy problem for their first return maps.

We thank the referee for his valuable assistance.

1. Preferred cross-sections

The most effective way to reduce a flow to a diffeomorphism would be a canonical cross-section or at least a finite number of preferred cross-sections. (From the viewpoint of the first-return maps, this is finding preferred elements in a flow equivalence class.) We will show in Theorem A that these exist for circular Axiom A flows [Sm] whose periodic orbits span $H_1(M;\mathbf{R})$.

We begin by observing that there is no reason to distinguish between cross-sections K_0 and K_1 to ρ if there is an isotopy of M carrying K_0 to K_1 through cross-sections, since K_0 and K_1 must determine conjugate first-return maps. Grouping such isotopic cross-sections together enables one to study cross-sections algebraically. Note that a cross-section K carries a preferred normal orientation, arising from the flow ρ, and so determines a dual class $u_K\in H^1(M;\mathbf{Z})$.

PROPOSITION [F2]. *There is an isotopy of M carrying K to L through cross-sections* $\Leftrightarrow u_K=u_L$.

This proposition suggests studying $\{u_K\in H^1(M;\mathbf{Z})\mid K \text{ is a cross section to } \rho\}$. For this purpose, topologize the *set of homology directions* D of M, $D=H_1(M;\mathbf{R})/\text{positive scalars}$, as (unit sphere)$\cup\{0\}$ and note that any closed loop γ

of M determines a homology direction $[\gamma] \in D$. Define $d \in D$ to be a *homology direction for* ρ if there are points $m_i \to m$ and times $t_i \to +\infty$ so that $\rho_{t_i} m_i \to m$ and $[(\rho_t m_i \mid 0 \le t \le t_i) \cdot \text{short path}] \to d$. Then the set D_ρ of homology directions for ρ is a compact subset of D.

THEOREM [F2]. *The smooth flow* ρ *on the compact manifold* M *has a cross-section* K *dual to* $u_K \in H^1(M; \mathbf{Z})$ *iff* $u_K(D_\rho) > 0$.

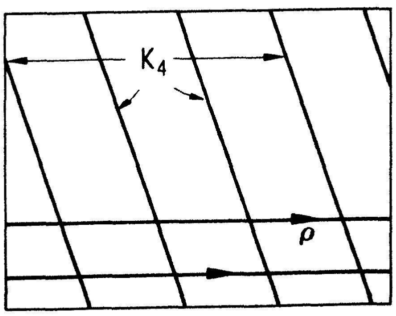
If ρ does have a cross-section and the first return map admits a Markov partition F of small size then there are a finite number of periodic orbits γ_i, $i = 1, \ldots, n$, corresponding to those allowed loops of elements of F which are minimal (no element of F occurs twice). One then has a simpler description of the cross-sections to ρ.

THEOREM [F2]. *For such* ρ, *a class* $u \in H^1(M; \mathbf{Z})$ *is dual to a cross-section* K *for* ρ *iff* $u(\gamma_i) > 0$, $i = 1, \ldots, n$.

Observe that one direction is easy, since $u_K(\gamma_i)$ is the number of times γ_i meets K in one period.

One should think of $u(D_\rho) > 0$ as defining an open cone $\mathscr{C} \subset H^1(M; \mathbf{R})$ in which lattice points (i.e., integral classes) correspond to cross-sections to ρ. (One may show that real classes in this open cone arise from closed 1-forms ω with $\omega(d/dt) > 0$). The flow ρ is circular exactly when $\mathscr{C}$ is nonempty (that is when D_ρ lies to one side of some hyperplane through the origin) in which case there are infinitely many distinct lattice points in $\mathscr{C}$. Under the hypotheses of the preceding theorem, $\mathscr{C}$ has finitely many flat, integrally defined sides. This fact will be exploited to find preferred lattice points in $\mathscr{C}$ and, in turn, preferred cross-sections.

These preferred cross-sections will be "simplest" in a certain sense. For instance, the flow $\rho_t(x, y) = (x, ye^{it})$ on $S^1 \times S^1$ has cross-sections $K_n = \{(z^n, z)\}$ for $n > 0$, with first return map f_n = rotation through $1/n$ of a revolution. Intuitively, K_n should be simple when n is small. This makes sense if one considers Haar



measure μ on $S^1 \times S^1$ and notes that μ induces an f_n-invariant measure

$$\bar{\mu}_n(S) = \frac{1}{\eta}\,\mu(\{\rho_t x \mid x \in K_n, 0 \le t \le \eta\})$$

(note that this quantity is independent of η if η is small). The total mass $\bar{\mu}_n(K_n) = n$ does increase with n. This captures intrinsically the idea that K_n is n times larger than K_1.

It is easy to generalize these considerations to any flow ρ with an invariant measure μ. If ρ has a cross-section K and K has first-return map f then $\bar{\mu}_K$ as defined above is independent of $\eta > 0$ (since μ is invariant) and gives a f-invariant, nonnormalized measure on K. The *total mass* $\bar{\mu}(K) = \bar{\mu}_K(K)$ measures the complexity of K from the viewpoint of the measure μ.

THEOREM A. *Suppose ρ is a smooth circular flow on a compact manifold M and there are classes $c_1, \ldots, c_n \in H_1(M; \mathbf{Z})$ so that*

1) *$u \in H^1(M; \mathbf{Z})$ is dual to a cross-section to ρ*

$$\Leftrightarrow u(c_i) > 0$$

2) *$\{c_i\}$ spans $H_1(M; \mathbf{R})$.*

Then there is a finite set $\mathcal{F} = \{K_1, \ldots, K_s\}$ of cross sections to ρ with the following properties:

a) *For any ρ-invariant measure μ there is a $K_i \in \mathcal{F}$ at which the total mass function $\bar{\mu}(K)$ achieves its minimum and the entropy $h_{\bar{\mu}}(f_K)$ achieves it maximum, over all cross-sections K to ρ and all return maps f_K.*

b) *Likewise the topological entropy $h(f_K)$ achieves its maximum over all cross-sections K to ρ at some $K_i \in \mathcal{F}$.*

Note. This applies, as promised, to circular Axiom A flows if the homology classes of closed orbits span $H_1(M; \mathbf{R})$.

Proof. By 1) the open cone $\mathscr{C}$ has finitely many flat, integrally defined sides. We want to define the K_i to be those cross-sections corresponding to the extreme points of the convex hull of the lattice points in $\mathscr{C}$. This will be satisfactory after showing:

LEMMA. *The convex hull of $\mathscr{C} \cap H^1(M; \mathbf{Z})$ has a finite positive number of extreme points.*

Proof. The semigroup $\{u \in H^1(M; \mathbf{Z}) \mid u(c_i) \ge 0, i = 1, \ldots, n\}$ is easily seen to have finitely many generators $u_1, \ldots, u_d$. If $u = \sum n_j u_j, 0 \le n_j \in \mathbf{Z}$, lies in

$\mathscr{C} \cap H^1(M;\mathbf{Z})$ and if some n_j (say n_1)>1, then $u \pm u_1 \in \mathscr{C} \cap H^1(M;\mathbf{Z})$ and so

$$u = \frac{u+u_1}{2} + \frac{u-u_1}{2}$$

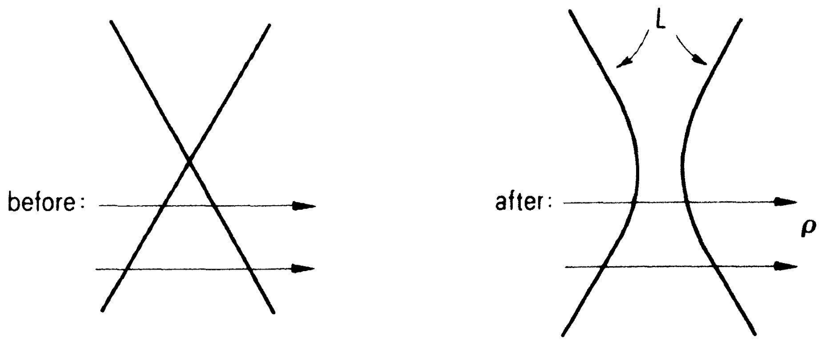
is not an extreme point of $\mathscr{C} \cap H^1(M;\mathbf{Z})$. Hence there are $\leq 2^d$ points.

Since ρ is circular, $\mathscr{C} \cap H^1(M;\mathbf{Z})$ is nonempty. Thus its convex hull is nonempty and has an extreme point.

Now suppose μ is some ρ-invariant measure. Then the total mass $\bar{\mu}$ deposited by μ on cross-sections is additive, in the following sense.

LEMMA. *If K_1 and K_2 are cross-sections to ρ and L is a cross section such that $u_L = u_{K_1} + u_{K_2}$ then $\bar{\mu}(L) = \bar{\mu}(K_1) + \bar{\mu}(K_2)$.*

Proof. One may assume that K_1 and K_2 intersect transversely and that L is the cross-section obtained by smoothing $K_1 \cap K_2$ in a product neighborhood as shown:



The uniqueness of cross-sections in a given cohomology class up to ambient isotopy implies that μ_K determines $\bar{\mu}(K)$. It is clear, however, that $\bar{\mu}(L) = \bar{\mu}(K_1) + \bar{\mu}(K_2)$. Q.E.D.

Since $\mathscr{C}$ is an open cone, $\bar{\mu}$ extends uniquely to a linear functional on $H^1(M;\mathbf{R})$ with positive values on $\mathscr{C}$. It follows that $\bar{\mu} \mid H^1(M;\mathbf{Z}) \cap \mathscr{C}$ is minimized on $\mathscr{F}$ (of course, this minimum may be assumed elsewhere as well).

By Abramov's formula [DGS],

$$h_\mu(\rho_1) = h_{\bar{\mu}}(f_K)\bar{\mu}(K).$$

So as K varies, $h_{\bar{\mu}}(f_K)$ is maximized on $\mathscr{F}$.

Finally, by Dinaburg-Goodwyn-Goodman [DGS], $h(f_K) = \sup_\nu h_\nu(f_K)$, where ν varies over the f_K-invariant measures. But such a ν gives rise to a ρ-invariant μ with $\bar{\mu} = \nu$. Hence $h(f_K) = \sup_\mu h_{\bar{\mu}}(f_K)$ is maximized on $\mathscr{F}$ as well. Q.E.D.

It should be noted that the "total mass" functional $\bar{\mu}$ used in the proof is the asymptotic cycle of μ [S]. Much of the above proof could be rephrased in Schwartzman's context, but the application to Axiom A systems requires homology directions.

One may naturally ask whether a generic set of flows with cross-sections possess homology classes satisfying 1) in Theorem A. This holds in dimension two thanks to the density of Morse-Smale diffeomorphisms on S^1. Such maps have rational rotation number which implies that the homology direction of the suspended flow consists of a single integral class c_1. The answer is not known in general.

The irrational flows on T^2 demonstrate the necessity for an integrality assumption such as 1). Here D_ρ is a single irrational vector, there is only one ρ invariant measure μ and $\bar{\mu}$ does not assume its mimimum value of 0. A cross section with very small mass may be obtained by taking a nearly closed flowline, joining its ends by a short transversal and then rounding the flowline to make it transverse to the flow. The mass deposited on this cross-section is essentially the length of the transversal.

Assumption 2) is somewhat less essential, as the following theorem shows.

THEOREM B. *Suppose ρ is a smooth flow on a compact manifold M and there are finitely many classes $c_i \in H_1(M;\mathbf{Z})$ such that ρ has a cross section dual to $u \Leftrightarrow u(c_i) > 0$, $i = 1, \ldots, n$. Then there are finitely many preferred families $F_1, \ldots, F_s$ of cross-sections, where two cross-sections are in the same family iff they take the same values on the c_i's. For any invariant measure μ, $\bar{\mu}$ is minimized and $h_{\bar{\mu}}(f_K)$ is maximized on $F_1 \cup \cdots \cup F_s$. Topological entropy is also maximized on $F_1 \cup \cdots \cup F_s$. The first-return maps for members of a given family have the same Artin-Mazur zeta function.*

Proof. By applying the arguments used in Theorem A to $H^1(M;\mathbf{Z})/\{u \mid u(c_i) = 0, i = 1, \ldots, n\}$ and $\{c \in H_1(M;\mathbf{Z}) \mid \text{for some } m > 0, mc = \sum m_i c_i\}$ (instead of $H^1(M;\mathbf{Z})$ and $H_1(M;\mathbf{Z})$) one again obtains the families $F_1, \ldots, F_s$ as the extreme lattice points in an open cone. The minimizing and maximizing properties follow just as before.

Assume the cross-sections K and L are in the same family, and let γ be a periodic orbit for ρ. If $[\gamma]$ were not expressible as a linear combination of c_i with nonnegative coefficients there would be an integral linear functional u with $u(c_i) > 0$ and $u(\gamma) < 0$. Such a u would be dual to a cross-section J with $u_J(\gamma) = u(\gamma) < 0$, which is impossible. Hence $[\gamma]$ is a linear combination of c_i's. As $u_K(c_i) = u_L(c_i)$ by our definition of family, it follows that $u_K(\gamma) = u_L(\gamma)$.

Clearly for given K and m the values of $u_K(\gamma)$ for all γ determine $N_m(f_K) = \#\mathrm{Fix}\,(f_K^m)$. Since the Artin-Mazur zeta function $\zeta(g) = \exp\left(\sum \frac{1}{m} N_m(g) t^m\right)$ depends only on $N_m(g)$, we have $\zeta(f_K)=\zeta(f_L)$ as desired. (If some $N_m(f_K)$ is infinite, ζ still makes sense formally as an indexing device for the coefficients.) Q.E.D.

One may say more if M is 2-dimensional. The only nontrivial case of Theorem B is when $M=T^2$, and ρ is a flow with cross-section whose first return map has rational rotation number. Then D_ρ is a single direction, represented by a unique indivisible $d\in H_1(M;\mathbf{Z})$/torsion. There is a single preferred family, namely $F_1=\{u\in H^1(M;\mathbf{Z})\mid u(d)=1\}$. It isn't hard to verify that F_1 consists of all the cross-sections K to ρ whose first return maps f_K have rotation number 0. Moreover, if $K, L\in F_1$, the maps f_K and f_L are conjugate; the flow ρ establishes a natural conjugacy between $\operatorname{Fix}(f_K)$ and $\operatorname{Fix}(f_L)$ and also a compatible sense of motion on the complementary intervals.

2. The homology zeta function for flows

The last conclusion of Theorem B suggests studying the way in which the Artin–Mazur zeta functions vary from one family to another. We will show that all these zeta functions are related to a new zeta function $\zeta_H(\rho)$ of several variables in a simple way. In the special case of the semiflow associated to the Lorenz attractor, the analogous zeta function was introduced by Williams [W].

For the sake of contrast, we will summarize some results about the usual zeta function for flows. Motivated by the Selberg zeta function for surfaces of constant negative curvature, Smale defined [Sm]

$$\zeta(s)=\prod_{\gamma,\,k}(1-e^{-(s+k)l(\gamma)})$$

where γ ranges over the prime periodic orbits of the smooth flow ρ and where $l(\gamma)$ is the period of γ. The product converges for $\operatorname{Re}(s)$ large provided the $l(\gamma)$ increase sufficiently rapidly. For constant-time suspensions of Axiom A diffeomorphisms and for geodesic flows on surfaces of constant negative curvature, $\zeta(s)$ is meromorphic in the whole plane [Sm]. For certain Axiom A flows, $\zeta(s)$ has an essential singularity at $s=-\varepsilon[G]$. By results of Bowen, the study of $\zeta(s)$ for Axiom A flows reduces to questions about the suspended flows of subshifts of finite type [B1].

Instead of monitoring the length of closed orbits, we will keep track of their homology class (modulo torsion). This homology zeta function $\zeta_H(\rho)$ for flows ρ is modeled on the Artin-Mazur zeta functions $\zeta(f)$ for diffeomorphisms f (as opposed to the Selberg zeta function mentioned above). The arguments of

Manning that establish the rationality of $\zeta(g)$ for Axiom A g carry over to give the rationality of $\zeta_H(\rho)$ for Axiom A flows ρ.

As with all zeta functions considered in dynamical systems, a finiteness assumption is needed to define $\zeta_H(\rho)$. Henceforth we assume ρ is a smooth flow on a compact manifold M with only finitely many periodic orbits in each homology class. (As we distinguish orbits from their multiples, this constraint excludes a flow with a closed orbit γ whose homology class $[\gamma]$ is of finite order in $H_1(M; \mathbf{Z})$.) We call such a flow ρ *homology finite*.

We now construct certain algebraic receptacles for the zeta functions to be defined. For a free abelian group G of finite rank, one may form the "group ring" $\mathbf{Z}[G] = \bigoplus_{g\in G} \mathbf{Z}g$ and the "rational formal power series module" $P_G = \prod_{g\in G} \mathbf{Q}g$. By regarding G as a multiplicative group, $\mathbf{Z}[G]$ has a natural multiplication and is an integral domain. Using the natural action of G on P_G by translation, P_G is a module over $\mathbf{Z}[G]$. In case rank $G = 1$, $G = \{t^i \mid i \in \mathbf{Z}\}$, one sees that $\mathbf{Z}[G]$ consists of finite integral Laurent series in t, that P_G consists of rational formal power series in t and t^{-1} and that the module action is multiplication. For rank $G > 1$, one obtains a similar interpretation for $\mathbf{Z}[G]$ and P_G upon choosing an integral basis for G.

DEFINITION. If ρ is homology finite, let

$$\log \zeta_H(\rho) = \sum_{\gamma} \sum_{k>0} \frac{[\gamma]^k}{k} \in P_H$$

where H is the finitely generated free abelian group $H_1(M; \mathbf{Z})/\text{torsion}$, γ varies over the prime periodic orbits of ρ and $[\gamma] \in H$ is the torsion-free part of the homology class of γ.

We check that $\log \zeta_H(\rho)$ is well-defined. As mentioned above, ρ homology finite implies $[\gamma] \neq 0$ for all γ. Since the torsion subgroup of $H_1(M; \mathbf{Z})$ is finite, only finitely many γ have the same $[\gamma]$. Together with the fact that any nonzero $h \in H$ is uniquely expressible as a positive power of an indivisible element, it follows that the coefficients in the formula are indeed finite and rational.

Note that $\zeta_H(\rho)$ is only implicitly defined via its logarithm. In the cases discussed below, $\zeta_H(\rho)$ exists in its own right. Note that for now that formally $\log \zeta_H(\rho) = \sum_\gamma -\log(1-[\gamma])$, that is $\zeta_H(\rho) = \prod_\gamma (1-[\gamma])^{-1}$.

The following is immediate from the definitions.

PROPOSITION 1. *If ρ is a flow with cross-section K dual to $u \in H^1(M; \mathbf{Z})$ then $\log \zeta(f_K) \in P_{\mathbf{Z}}$ is obtained from $\log \zeta_H(\rho)$ by substituting $u[\gamma]$ for $[\gamma]$.* (*Here one thinks of $\mathbf{Z}$ as the free multiplicative group on one generator t.*)

Unfortunately the module P_H has nontrivial torsion (as shown the author by W. Dwyer). This gives rise to ambiguities in interpreting a quotient of elements of $\mathbf{Z}[H]$ as elements of P_H. For instance, $1/t+t^{-1}$ may be expanded as either $t-t^3+t^5-t^7+\cdots$ or as $t^{-1}-t^{-3}+t^{-5}-t^{-7}+\ldots$. We may avoid these ambiguities by restricting the quotients we consider as in the following definition.

DEFINITION. The formal power series $s\in P_G$ is *strictly rational* if there are $p, q\in \mathbf{Z}[G]$ and $a, b\in \mathbf{Z}$, $b\neq 0$ such that

1) regarding G additively, 0 is not a convex combination of the terms in p, q with nonvanishing coefficients and

$$2)\ s=(a+p)\frac{1}{b}\left(1-\frac{q}{b}+\frac{q^2}{b^2}-\frac{q^3}{b^3}+\cdots\right)$$

in which case one writes $s=(a+p)/(b+q)$. Note that condition 1) is equivalent to the existence of a linear functional $u: G\to \mathbf{Z}$ positive on the terms of p and g. The infinite series consequently contributes only finitely many coefficients to a given $g\in G$ (it is a "locally finite" sum) and hence defines an element of P_G. It is easy to check, regarding $p\in \mathbf{Z}[G]$ as an element of P_G in the obvious way, that is $(a+p)=(b+q)s$.

If s is strictly rational as just defined and if $a=b$, then one may define $\log s$ using the power series for $\log 1+x$:

$$\log s = x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots$$

where $x=a^{-1}p+(1+a^{-1}p)(-a^{-1}q+a^{-2}q^2-a^{-3}q^3+\cdots)$. As before, one may use the functional u to see that this is a locally finite sum, and hence gives a well-defined element of P_G.

DEFINITION. If the power series $s_1,\ldots, s_k$ are strictly rational and $a_i=b_i$ for $i=1,\ldots,k$ then the power series $\log s_1+\cdots+\log s_k$ will be called the *formal logarithm of the rational function* $\prod_{i=1}^k s_i$. One must be careful not to manipulate the formal product $\prod_{i=1}^k s_i$ algebraically. For instance the rational function

$$\frac{1}{1+t^2}\cdot\frac{1}{1+t^{-1}}$$

has formal logarithm

$$\log\frac{1}{1+t^2}+\log\frac{1}{1+t^{-1}}=\left(\cdots\frac{t^{-2}}{2}-t^{-1}-t^2+\frac{t^4}{2}-\frac{t^6}{3}+\cdots\right)\in P_{\mathbf{Z}}.$$

But

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

cannot be expanded as $1-(t^{-1}+t+t^2)+(t^{-1}+t+t^2)^2-\cdots$ since the coefficients are all given by divergent series.

We shall prove

THEOREM C. *For any Axiom A homology finite flow ρ, $\zeta_H(\rho)$ is rational.*

Note that this means precisely that there are strictly rational function s_i with $a_i=b_i$, $i=1,\ldots,k$, so that the series $\log\zeta_H(\rho)$ defined above equals $\sum_{i=1}^k \log s_i$. In fact, we will use one s_i for each basic set Γ for ρ. Define $\log\zeta_H(\rho\mid\Lambda)$ to be the element of P_H obtained by restricting the orbits γ which appear in the definition of $\log\zeta_H(\rho)$ to those orbits $\gamma\subset\Lambda$. The theorem clearly will follow from the following proposition, by Smale's Spectral Decomposition Theorem [Sm].

PROPOSITION 2. *Given a basic set Λ for ρ, there is a strictly rational function s with $a=b=1$ such that*

$$\log\zeta_H(\rho\mid\Lambda)=\log s.$$

Proof of the proposition. Recall that Λ admits a Markov family of sections $\mathcal{M}$ and let $\gamma_1,\ldots,\gamma_m$ be the closed orbits determined by minimal allowed loops $l_1,\ldots,l_m$ of elements of F. Suppose that some convex combination of $[\gamma_1],\ldots,[\gamma_m]\in H$ gives 0. Then for some nonnegative integers $a_1,\ldots,a_m$, one has $\sum a_i[\gamma_i]=0$. Choose $x_0\in\mathcal{M}$ and $x_i\in l_i$ for $i=1,\ldots,m$. There are allowed sequences p_i from x_0 to x_i and q_i from x_i to x_0. For any $n\geq 0$, the closed orbit σ_n determined by

$$p_1\circ\underbrace{l_1\cdots l_1}_{na_1}\circ q_1\circ p_2\circ\underbrace{l_2\cdots l_2}_{na_2}\circ q_2\circ\cdots\circ q_n$$

satisfies $[\delta_n]=[\delta_0]$. Since the map carrying symbol sequences to orbits is finite to one, the $\{\delta_n\}$ would be an infinite collection of orbits corresponding to the same class $[\delta_0]\in H$, contradicting the homology finiteness of ρ.

Since 0 is not a convex combination of $[\gamma_1],\ldots,[\gamma_m]$ it follows that there exists $u\in H^1(M;\mathbf{Z})$ with $u[\gamma_i]>0$, all i. By [F2] there is a map $\theta: M\to S^1$ so that $d(\theta\circ\rho_t)/dt>0$, near Λ and $\theta^*(1)=u$. Roughly speaking, there is a cross-section near Λ in class u. It follows easily from this that for any $n>0$ only finitely many closed orbits $\gamma\subset\Lambda$ satisfy $u(\gamma)=n$.

Hence $\log\zeta_H(\rho\mid\Lambda)\in P_H(u)=\{x\in P_H\mid$ only finitely many terms c_h, $c_h\neq 0$, in x satisfy $u(h)=n$, any $n>0$ and no terms have $u(h)\leq 0$ and $h\neq 0\}$. It is easy to check that $P_H(u)$ is an integral domain. The units of $P_H(u)$ are precisely $\{x=\sum c_h h\in P_H(u)\mid c_0\neq 0\}$.

For the rest of this proof, all our computations shall be made in this integral domain.

Our approach will follow Bowen's scheme for reducing properties of closed orbits of $\rho\mid\Lambda$ to symbolic dynamics using Markov partitions [B1]. Bowen (using Manning's work with diffeomorphisms as a guide) constructs a finite index set I, using a function $l(i): I\to\mathbf{Z}^+$ such that for each $i\in I$ one has the following gadgetry:

1) A subshift of finite type $\sigma_i: Y_i\to Y_i$
2) A continuous function $t_i: Y_i\to(0,\infty)$
3) The suspended flow ψ_i on $X_i=\{(y,t)\mid y\in Y_i,\ 0\leq t\leq t_i(y)\}/(y,t_i(y))=(\sigma_i(y),0)$ and
4) A continuous, finite-to-one map $\pi_i: X_i\to\Omega(\rho)$ such that $\rho_t\circ\pi_i=\pi_i\circ\psi_{i,t}$ for all $t\in\mathbf{R}$.

The key property of this set-up is the relation, for each periodic point $p\in\Lambda$ and $\tau\in\mathbf{Z}^+$,

$$\sum_{i\in I}(-1)^{l(i)+1}\operatorname{card}\{x\mid\pi_i(x)=p,\ \tau=\text{prime period of } x)$$
$$=\begin{cases}1 & \text{if } \tau \text{ is the prime period of } p\\ 0 & \text{otherwise}\end{cases}$$

In our context, this yields

$$\log\zeta_H(\rho)=\sum_{\gamma\subset\Lambda}\sum_{k>0}\frac{[\gamma]^k}{k}=\sum_{i\in I}(-1)^{l(i)+1}\left(\sum_{\delta}\sum_{k>0}\frac{[\pi_i(\delta)]^k}{k}\right)$$

where δ varies over the prime periodic orbits for ψ_i.

The Wang sequence of the fibration $Y_i \to X_i \to S^1$ in Cech cohomology

$$H^0(Y_i;\mathbf{Z}) \xrightarrow{f_i^*-1} H^0(Y_i;\mathbf{Z}) \xrightarrow{d} H^1(X_i;\mathbf{Z}) \to 0$$

is exact. Note that $H^0(Y_i;\mathbf{Z}) = \{$continuous maps from Y_i to $\mathbf{Z}\}$ is generated by the characteristic functions of closed-open sets. Any closed-open set in Y_i is the union of cylinder sets for some Markov partition of Y_i (simply perform symbol splitting enough times on a given Markov partition). Because $H^1(M;\mathbf{Z})$ is free and d is surjective, $\pi_i^*: H^1(M;\mathbf{Z}) \to H^1(X_i;\mathbf{Z})$ lifts to $H^0(Y_i;\mathbf{Z})$. The facts just mentioned show how to find a homomorphism L such that

$$\begin{tikzcd} H^1(M;\mathbf{Z}) \arrow[r, "L"] \arrow[dr, "\pi_i^*"'] & F \subset H^0(Y_i;\mathbf{Z}) \arrow[d, "d"] \\ & H^1(X_1;\mathbf{Z}) \arrow[d] \\ & 0 \end{tikzcd}$$

commutes. Here, for some Markov partition $\mathcal{M}$ for σ_i, $F = \mathrm{Cyl}(\mathcal{M}) = \{f: Y_i \to \mathbf{Z} \mid f$ is constant on each cylinder set $C \in \mathcal{M}\}$. Note that F is a free abelian group, naturally dual to the free abelian group on $\mathcal{M}$. Let A_i be the 0-1 matrix of transitions for $\mathcal{M}$.

Let ϵ vary over all sequences $(C_{j_1}, \ldots, C_{j_p})$ of cylinder sets in $\mathcal{M}$ satisfying $A_i(C_{j_k}, C_{j_{k+1}}) = 1$, k modulo p. Such a sequence ϵ has a minimal period $\frac{p}{k}$, repeated k times. Also, ϵ determines the σ_i-periodic orbit $\bar{\epsilon}$ that passes cyclically through $C_{j_1}, \ldots, C_{j_p}$ and a given $\bar{\epsilon}$ appears for precisely the $\frac{p}{k}$ cyclic permutations of ϵ.

This gives

$$\sum_{\delta}\sum_{k>0} \frac{[\pi_i(\delta)]^k}{k} = \sum_{\epsilon} \frac{[\pi_i(\bar{\epsilon})]}{p}$$

By taking the product of all the cylinder sets over one period, a σ_i-periodic orbit $\bar{\epsilon}$ determines an element $[\bar{\epsilon}] \in F^* = \mathrm{Hom}(F;\mathbf{Z})$ (recall that F^* is the free abelian group on $\mathcal{M}$). Using the above diagram of maps we find

$$\sum_{\epsilon} \frac{[\pi_i(\bar{\epsilon})]}{p} \in P_H$$

is obtained from

$$\sum_{\epsilon} \frac{[\bar{\epsilon}]}{p} \in P_{F^*}$$

by replacing each term af^* by $a \cdot L^T f^*$, where $L^T: F^* \to (H^1(M; \mathbf{Z}))^* = H$.

We now compute

$$\begin{aligned}
\sum_{\epsilon} \frac{[\bar{\epsilon}]}{p} &= \sum_{\epsilon} \frac{C_{j_1} \cdots C_{j_p}}{p} \\
&= \sum_{p>0} \frac{1}{p} \operatorname{Trace}(CA_i)^p, \quad \text{where } C = \begin{pmatrix} C_1 & & 0 \\ & \ddots & \\ 0 & & C_d \end{pmatrix}, \ d = \operatorname{card} \mathcal{M} \\
&= -\log \det (I - CA_i)
\end{aligned}$$

Applying L^T term by term gives

$$\sum_{\epsilon} \frac{[\pi_i(\bar{\epsilon})]}{p} = -\log \det (I - B_i),$$

where the entries of B_i are either 0 or in H. Using the fact that $u(L^T(C_{j_1} \cdots C_{j_p})) = Lu(C_j, \ldots, C_{j_p}) = u[\pi_i(\bar{\epsilon})] > 0$ for all ϵ, we see that $-\log \det (I - B_i) \in P_H(u)$.

Finally, we have

$$\log \zeta_H(\rho) = \sum_{i \in I} (-1)^{l(i)+1} (-\log \det (I - B_i)).$$

Setting

$$1 + p = \prod_{l(i) \text{ even}} \det (I - B_i) \quad \text{and} \quad 1 + q = \prod_{l(i) \text{ odd}} \det (I - B_i)$$

yields $\zeta_H(\rho) = (1+p)/(1+q)$, so we are done. Q.E.D.

COROLLARY. *If ρ is an Axiom A flow with cross-section, all the zeta functions of first-return maps are obtained by substituting monomials $x_1 = t^{a_1}, \ldots, x_\beta = t^{a_\beta}$ into a rational function $\zeta_H(\rho)$ of $x_1, \ldots, x_\beta$. Here $\beta = \operatorname{rank} H_1(M; \mathbf{Z})$ and $(a_1, \ldots, a_\beta) \in \mathbf{Z}^\beta$ are the coordinates of the class in $H^1(M; \mathbf{Z})$ dual to the cross-section.*

Proof. This is almost immediate. One should note, however, that since ρ has a cross-section there is a functional $u: H \to \mathbf{Z}$ positive on all classes of closed orbits of ρ, and this functional should be used in the proof of Proposition 2 for all the basic sets of ρ. Once this precaution is taken, $\zeta_H(\rho)$ may be unambiguously interpreted as the result of applying the exponential series $\sum x^n/n!$ formally to $\log \zeta_H(\rho)$, since all sums will be locally finite. Q.E.D.

3. The homology zeta function and basic sets

When Λ is a basic set for an Axiom A flow ρ and Λ contains only finitely many periodic orbits in each homology class, the proof of Proposition 2, section 2 shows that $\zeta_H(\rho \mid \Lambda)$ is well-defined and lies in the quotient field $Q(\mathbf{Z}[H])$ of $\mathbf{Z}[H]$. We here study $\zeta_H(\rho \mid \Lambda)$ as an invariant of the basic set Λ. The homology finiteness assumption for $\rho \mid \Lambda$ will be interpreted dynamically. It will also be shown that $\zeta_H(\rho \mid \Lambda)$ usually determines $D_{\rho|\Lambda}$, the homology directions of ρ on Λ (although in general one extra piece of information is needed). This leads to a canonical choice of the classes c_i referred to in Theorem A above.

THEOREM D. *For a basic set Λ of an Axiom A flow ρ on a compact manifold M, the following are equivalent:*
1) *$\rho|\Lambda$ is homology finite, i.e., only finitely many orbits of $\rho|\Lambda$ lie in any given class in $H_1(M; \mathbf{Z})$.*
2) *Some $u \in H^1(M; \mathbf{Z})$ is positive on all closed orbits of $\rho|\Lambda$.*
3) *Some $u \in H^1(M; \mathbf{Z})$ is positive on $D_{\rho|\Lambda}$.*
4) *There is a compact, codimension one submanifold $K \subset M$ with a preferred normal orientation so that each flowline in Λ meets K and all such intersections are transverse and in the positive sense.*

If these conditions hold, the proof of Proposition 2, Section 2 yields the formula $\zeta_H(\rho) = (1+p)/(1+q)$, where u is positive on all the terms $\{h_i\} \subset H$ of $p, q \in \mathbf{Z}[H]$. The convex hull $C \subset H \otimes \mathbf{R} = H_1(M; \mathbf{R})$ of $\{h_i\}$ does not contain 0, and the projection of C to $D = H_1(M; \mathbf{R})$/positive scalars equals $D_{\rho|\Lambda}$.

If one requires $1+p$ and $1+q$ to be relatively prime, then p and q are uniquely determined, independently of the choice of Markov family and the functional u. The same characterization of $D_{\rho|\Lambda}$ holds.

Proof. 1) $\to$ 2) was already shown in Proposition 2, Section 2

2) $\to$ 3) follows from [F2]. When 2) holds, $D_{\rho|\Lambda}$ is the projection of D of the convex hull of the classes of closed orbits (indeed, a finite number of classes suffice).

3) $\to$ 4) also follows from [F2]. When 3) holds, there is a function $\theta: M \to S^1$ with $\theta^*(1) = u$ and $(d/dt)\theta(\rho_t x) > 0$. By choosing θ smooth, one may let K be the inverse image of a regular value of θ.

4) $\to$ 1) since there is a continuous first return map on $K \cap \Lambda$ which, by hyperbolicity, has isolated periodic points of given period. For $n > 0$, the compactness of $K \subset \Lambda$ implies only finitely many closed orbits $\gamma \subset \Lambda$ satisfy $u_K(\gamma) = n$, and 1) follows.

Writing $p = \sum m_i h_i$, $q = \sum n_i h_i$, where for each i at least one of m_i, n_i is nonzero, one has (from the proof of Proposition 2, Section 2) that any $u \in H^1(M; \mathbf{Z})$ which is positive on the closed orbits of $\rho|\Lambda$ is also positive on $\{h_i\}$. Conversely, if u is positive on $\{h_i\}$ then u is positive on the terms of $\log \zeta_H(\rho \mid \Lambda) = \log(1+p) - \log(1+q)$. Since, as mentioned above, $D_{\rho|\Lambda}$ is the projection to D of the convex hull of the homology classes of closed orbits, one obtains the desired description of $D_{\rho|\Lambda}$.

To obtain the canonical choice of p, q, one uses the unique factorization property of $\mathbf{Z}[H]$. The intrinsic definition of $\zeta_H(\rho \mid \Lambda)$ shows that the ratio $(1+p)/(1+q) \in Q(\mathbf{Z}[H])$ is independent of Markov partition. A complete factorization of $1+p$ or $1+q$ into irreducibles may be modified factor by factor with units $\pm h$, $h \in H$ so that each factor is of form $1+r$, where u is positive on the terms of r. These factorizations of $1+p$, $1+q$ are unique up to order. By eliminating common factors $1+r$, one obtains the canonical expression $\zeta_H(\rho \mid \Lambda) = (1+p')/(1+q')$. By the proof of Proposition 2, Section 2, one has that p, q (hence p', q') are independent of the choice of u. Arguing as in the last paragraph, we are done. Q.E.D.

If for each basic set Λ for ρ the terms h_i are chosen from the canonical choice of p, q, then $\bigcup_\Lambda \{h_i\}$ is a canonical choice of $\{c_i\}$ for Theorem III above. For by [F2], $u(D_\rho) > 0 \Leftrightarrow u(D_{\rho|\Lambda}) > 0$ for all Λ.

Suppose one wishes to compute $D_{\rho|\Lambda}$ from $\zeta_H(\rho \mid \Lambda) \in Q(\mathbf{Z}[H])$. If one knows a functional $u: H \to \mathbf{Z}$ positive on $D_{\rho|\Lambda}$ then the preceding proof gives a procedure for calculating $D_{\rho|\Lambda}$ via a canonical form for $\zeta_H(\rho \mid \Lambda)$. It is often unnecessary, however, to know such a functional. Express $\zeta_H(\rho \mid \Lambda)$ in lowest terms as r/s. Then, $r, s \in \mathbf{Z}[H]$ are determined up to units, i.e., can only be replaced by $\epsilon r, \epsilon s$ with $\epsilon = \pm h$, $h \in H$. One may look at the convex hulls $C_r, C_s \subset H \otimes \mathbf{R}$ of the terms appearing in r, s. There must be one point h which is an extreme point both for C_r and C_s and at which the coefficients r_h, s_h satisfy $r_h = s_h = \pm 1$. Usually there will be only one such point h and the canonical form of $\zeta_H(\rho \mid \Lambda)$ will be

$$\frac{(r_h h)^{-1} r}{(s_h h)^{-1} s} = \frac{1+p}{1+q}.$$

In any case, p and q (and hence $D_{\rho|\Lambda}$) are determined up to finite ambiguity by $\zeta_H(\rho \mid \Lambda)$.

4. Flow equivalence and topological entropy

Theorems A and B indicate where to look in a flow equivalence class for the maximum value of the topological entropy, under certain rationality assumptions on the homology directions. In this section we will further investigate how topological entropy varies under flow equivalence.

With $\mathscr{C} \subset H^1(M; \mathbf{R})$ as in Section 1, each lattice point $u \in \mathscr{C} \cap H^1(M; \mathbf{Z})$ gives rise to a cross-section K, a return map r (determined by u up to conjugacy) and a topological entropy $h(u) = h(r)$. For any $n > 0$, the cross-section L corresponding to $n \cdot u$ is n disjoint copies of K, cyclically permuted by the return map $s: L \to L$ in such a way that s^n consists of n disjoint copies of r. It follows that $h(nu) = (1/n)h(u)$, that is the function $h: \mathscr{C} \cap H^1(M; \mathbf{Z}) \to \mathbf{R}$ is homogeneous of degree -1. We will show

THEOREM E. *There exists a unique extension of h to a continuous map $\hat{h}: \mathscr{C} \to \mathbf{R}$ that is homogeneous of degree -1. Either $\hat{h}(\mathscr{C}) = 0$ or $\hat{h}(\mathscr{C}) \subset (0, \infty)$. When $\hat{h} > 0$ on $\mathscr{C}$, $1/h$ is concave.*

Proof. Suppose that for some $u \in H^1(M; \mathbf{Z}) \cap \mathscr{C}$ we have $h(r) = 0$. Then the Dinaburg-Goodwyn-Goodman Theorem [B2] gives $h_{\bar{\mu}}(r) = 0$ for all ρ-invariant measures μ. By Abramov's formula [DGS], $h_\mu(\rho_1) = 0$. Choosing another section and reversing this reasoning shows h vanishes on all of $H^1(M; \mathbf{Z}) \cap \mathscr{C}$. By homogeneity and continuity, $\hat{h}(\mathscr{C}) = 0$.

We may assume now that $\boldsymbol{h}$ is positive on $H^1(M; \mathbf{Z}) \cap \mathscr{C}$. As μ varies over those ρ invariant measures for which $h_\mu(\rho_1) > 0$, Dinaburg-Goodwyn-Goodman gives $h(r) = \sup_\mu h_{\bar{\mu}}(r)$. Abramov gives

$$\frac{1}{h(r)} = \inf_\mu \frac{\bar{\mu}(K)}{h_\mu \rho_1}.$$

Regarding $\bar{\mu}(K)$ as a linear functional positive on $\mathscr{C}$, the right hand side defines a nonnegative concave function $g:\bar{\mathscr{C}}\to\mathbf{R}$ that is homogeneous of degree $+1$. From concavity, we see g is continuous. As $g>0$ on the dense set $H^1(M;\mathbf{Q})\cap\mathscr{C}$, the concavity of g gives $g>0$ on $\mathscr{C}$. So $\hat{h}=1/g:\mathscr{C}\to\mathbf{R}$ is the desired extension. Q.E.D.

We may say more about $\hat{h}$ when ρ is hyperbolic.

THEOREM F. *Suppose ρ is the suspension flow of $f:J\to J$ where f is*

a) *an Axiom A diffeomorphism with perfect Ω or*
b) *a pseudo-Anosov map [FLP].*

Then $\hat{h}$ tends to ∞ on $\partial\mathscr{C}$.

Remark. Note that Theorem A shows that there is a finite maximum for $\hat{h}$ on the set of integral points in $\mathscr{C}$. To fix ideas, we present a simple function with the qualitative features of $\hat{h}$. Let $\mathscr{C}$ be the open positive quadrant in the $x-y$ plane and let $f(x, y)=x^2+y^2/xy(x+y)$. Then $1/f$ is homogeneous, concave, positive on $\mathscr{C}$ and vanishes on $\partial\mathscr{C}$ but $f(x, y)\leq 1$ for all positive integers x and y.

Proof. We will restrict ourselves to case a), as case b) is nearly identical.

Let Λ be a basic set for ρ. Suppose $u\in H^1(M;\mathbf{Z})$ is positive on $D_{\rho|\Lambda}$. Then Theorem D shows there is a submanifold K dual to u and transverse to $\rho\mid\Lambda$. There is an associated return map $r:K\cap\Lambda \circlearrowleft$ which is an infinite basic set. The associated entropy $h_\Lambda(u)$ is therefore positive [B3].

As in the proof of Theorem E one shows that $g_\Lambda=1/h_\Lambda$ extends uniquely to a concave function on $\bar{\mathscr{C}}_\Lambda$, where $\mathscr{C}_\Lambda=\{u\mid u(D_{\rho|\Lambda})>0\}$. As $\mathscr{C}=\bigcap_\Lambda\mathscr{C}_\Lambda$ and $h=\sup_\Lambda(h_\Lambda\mid\mathscr{C})$ we need only show g_Λ vanishes on $\partial\mathscr{C}_\Lambda$.

When $D_{\rho|\Lambda}$ consists of a single point, then Theorem B shows that the Artin-Mazur zeta function $\zeta(r)$ associated to $u\in\mathscr{C}_\Lambda$ depends only on $u(d)$. As the entropy of an Axiom A basic set is the growth rate of the number of periodic points, $h_\Lambda(u)$ depends only on $u(d)$. By homogeneity, $g_\Lambda(u)$ is proportional to $u(d)$ and tends to zero on $\partial\mathscr{C}_\Lambda$.

When D_ρ contains more than one point we must proceed differently. For simplicity, assume that Λ is the suspension flow of a subshift of finite type. Choose a class $v\in H^1(M;\mathbf{Z})$ with $v(D_{\rho|\Lambda})\geq 0$, $v(D_{\rho|\Lambda})\neq 0$, and $0\in v(D_{\rho|\Lambda})$. Let F be a Markov family of local sections to $\rho\mid\Lambda$. As stated in section 1, there is a minimal loop in F such that the associated orbit γ has $v(\gamma)=0$. Let $s\in F$ be one of the symbols occurring in γ. Then there is a loop beginning and ending at s such that the associated orbit δ has $v(\delta)>0$. By concatenating these loops at s one obtains a family of periodic orbits with zeta function $1/1-[\gamma]-[\delta]$.

Choose $u\in\mathscr{C}\cap H^1(M;\mathbf{Z})$. For all $n>0$, $u+nv\in\mathscr{C}\cap H^1(M;\mathbf{Z})$ so there is a corresponding return map r_n and zeta function $\zeta_n=\zeta(r_n)$. Recall that $h(r_n)$ is the growth rate of the number of periodic orbits of r_n. We estimate this from below using the orbits constructed in the preceding paragraph: $h(r_n)\geq h_n$ where e^{-h_n} is the smallest zero of $1-t^{(u+nv)(\gamma)}-t^{(u+nv)(\delta)}$. We let $a=u(\gamma)$, $b=u(\delta)$, $c=v(\delta)$ and $p(t,x)=1-t^a-t^bx^c$. Graphing $p(t,x)=0$ shows that this curve passes through $(t,x)=(1,0)$ and has a continuation from this point into $t<1$, $x>0$. Thus for n large there is an intersection point (t_n,x_n) of $p(t,x)=0$ with $x=t^n$ lying near $(1,0)$. Thus $e^{-nh_n}=x_n\to 0$, so $nh_\Lambda(u+nv)\geq nh_n\to\infty$. We get

$$g_\Lambda(v)=\lim_{n\to\infty} g_\Lambda\left(v+\frac{u}{n}\right)=\lim_{n\to\infty}\frac{1}{nh_\Lambda(u+nv)}=0.$$

By homogeneity, g vanishes on all rational points in $\partial\mathscr{C}_\Lambda$ not in $\operatorname{Ann}(D_{\rho|\Lambda})$. As such points are dense, g_Λ vanishes on $\partial\mathscr{C}_\Lambda$.

When Λ isn't the suspension of a subshift of finite type, one passes to the 1-dimensional flow ψ defined by a Markov partition. One shows in the usual way that $h(u)$ is also the entropy of the return map for ψ associated to u, and then the preceding argument applies. Q.E.D.

5. Flow equivalence of Anosov diffeomorphisms

Two interesting homological results are known for an Anosov flow ρ with cross-section on a compact connected manifold M. The optimistic result is that when ρ has a smooth invariant measure the homology classes of closed orbits span $H_1(M;\mathbf{R})$ [P]. We extend this at the end of this section to transitive ρ. Thus Theorem A applies for such ρ. As any known ρ is conjugate to one with a smooth invariant measure, our Theorem A applies to all known ρ. The pessimistic result is that for all these known cases $H^1(M;\mathbf{Z})=\mathbf{Z}$, hence ρ has only one connected cross-section and Theorem A is trivial for such ρ. One may try, however, to use flow equivalence to prove $H^1(M;\mathbf{Z})=\mathbf{Z}$ is general (i.e., not just for known cases). We here present the results of such an attempt, and so relate in an unobvious way several properties which are often conjectured to hold for all Anosov diffeomorphisms $A:K\to K$, K connected and compact. Some results of this section are in the author's thesis [F1] which was written under the direction of Stephen Smale.

One property is purely homological, namely

Property 1) No eigenvalue of $A^*:H^1(K;\mathbf{R})\to H^1(K;\mathbf{R})$ is a root of unity.

This holds for all known examples [Fr]; the most general known result [H] is that 1) is true if $\pi_1 K$ has a polycyclic subgroup of finite index and if the Betti numbers of the universal cover of K are finite. As will be shown later, the condition $H^1(M;\mathbf{R})=\mathbf{R}$ is equivalent to 1 not being an eigenvalue of A^*.

Another property is purely dynamical, namely

Property 2) A has a fixed point.

This holds for all known A, since they are conjugate to infranilmanifold examples for which 2) holds.

The last property is a homological property of the Anosov manifold K.

Property 3) $\dim H^*(K;\mathbf{R})\leq 2^k$, $k=\dim K$.

To check this for known examples, one may use finite covers and the transfer map to reduce to studying nilmanifolds K. By inducting on the nilpotent degree of K, and using the Serre spectral sequence for the fibration of K over a torus as in [M1], one obtains $\dim H^*(K;\mathbf{R})\leq\Pi\dim H^*(T_i;\mathbf{R})=2^k$, where T_i are the tori which appear as successive fibers in the representation of K as an iterated torus bundle. This line of argument was suggested to the author by Rufus Bowen.

We now state

THEOREM G. *If an Anosov diffeomorphism $A_0:K_0\to K_0$, K_0 compact and connected, fails to satisfy Property* 1) *then there is an Anosov diffeomorphism $A:K\to K$, with K compact and connected and $\dim K=\dim K_0$, so that*

a) *1 is an eigenvalue of $A^*:H^1(K;\mathbf{R})\to H^1(K;\mathbf{R})$*

b) *A has no periodic points up to any given period, and*

c) *$\dim H^*(K;\mathbf{R})$ is arbitrarily large.*

So if Property 2 (*or Property* 3) *holds for all Anosov diffeomorphism of a fixed dimension, Property* 1 *must hold as well.*

Proof. By passing to a finite cover, one may assume that the stable and unstable bundles of A_0 are orientable, as in [M]. By passing to a power of A_0, one may assume A_0 preserves these orientations and that 1 is an eigenvalue for $A_0^*:H^1(K;\mathbf{R})\to H^1(K;\mathbf{R})$. Suspending A_0 gives an Anosov flow ρ on a compact manifold M, $\dim M=\dim K_0+1$. From the Wang exact sequence on real cohomology, one has $0\to H^0K_0\to H^1M\to H^1K_0\xrightarrow{A^*-1}H^1K_0$, and so $\dim H^1(M;\mathbf{R})=\operatorname{rank} H^1(M;\mathbf{Z})>1$. Thus the flow equivalence class of A_0 is nontrivial; we shall choose A from this flow equivalence class.

Since ρ is Anosov, it satisfies Axiom A. Hence for a finite number of classes $c_i\in H$, $u\in H^1(M;\mathbf{Z})$ is dual to a cross-section to ρ iff $u(c_i)>0$ for all i. By adding some extra c_i if necessary, we may assume all convex combinations in H of $(0, c_i)$ are also c_i's.

Let u_0 be the indivisible cohomology class $p^*(1)$ where $p: M \to S^1$ is the natural fibration with fiber K_0 and $p^*: H^1(S^1; \mathbf{Z}) = \mathbf{Z} \to H^1(M; \mathbf{Z})$. Extend u_0 to an integral basis for $H^1(M; \mathbf{Z})$ and let u_1 be another vector in this basis. Then for $a = a(n)$ large enough, the class $u = u_1 + au_0$ satisfies:

1) $u(D_\rho) > 0$ and u is indivisible
2) $u(c_i) \geq n$ for all i.

By 2) and the choice of c_i's one has $u[\gamma] \geq n$ for all periodic orbits γ of ρ, proving part b) of the Theorem.

Let K be the cross-section dual to u and A the first return map of K. Then A is Anosov and has no periodic points of period $<n$. Consequently the Lefschetz number $L(A^i) = 0$ for $i = 1, \ldots, n-1$.

Were $\dim H^*(K; \mathbf{R}) < n$, it would follow from the algebraic argument in [Fu] that $L(A^i) = 0$ for all i. Since A is Anosov and preserves the unstable orientation, one has [Sm] the equality $|L(A^i)| = \#\mathrm{Fix}\,(A^i)$, $i > 0$. Since A has periodic points, this is a contradiction. Thus $\dim H^*(K; \mathbf{R}) \geq n$, proving part c) of the Theorem.

The Wang sequence for $K \to M \to S^1$ shows that 1 is an eigenvalue of A^*. This gives part a). Q.E.D.

The preceding arguments work just as well for Thurston's pseudo-Anosov maps, since flow-equivalence preserves pseudo-Anosov [FLP, Exp. 14]. However it is known [FLP, Exp. 13] that property 1 fails for certain pseudo-Anosov maps; indeed, the induced map on H_1 can even be the identity. This gives

COROLLARY. *Given any integer n there is a pseudo-Anosov map of a closed connected surface with no periodic points of period* $<n$.

The pseudo-Anosov case of the following result was first proven by Thurston [FLP, Exp. 14] by completely different methods (using that the mapping torus of a pseudo-Anosov map is atoroidal). The next theorem settles the problem Plante raised in [P] to generalize his result to all transitive Anosov flows from the volume preserving case. Plante used asymptotic cycles whereas we use $\mathbf{Z}$-covers as in [F2].

THEOREM H. *Let* $\rho_t: M \to M$ *be either a transitive Anosov flow or a pseudo-Anosov flow* (*that is, a flow with cross-section with a pseudo-Anosov return map*). *Then the homology classes of closed orbits span* $H_1(M; \mathbf{R})$.

Proof. Clearly we may take M connected.

Let $\mathcal{M}$ be a fine Markov partition for ρ and let $\gamma_1, \ldots, \gamma_k$ be the closed ρ orbits corresponding to minimal loops in $\mathcal{M}$ [F2]. As in [F2, Theorem H], the homology class $[\gamma]$ of any periodic orbit γ is a positive integral combination of $[\gamma_1], \ldots, [\gamma_k]$. Assuming these classes don't span $H_1(M; \mathbf{R})$, then some indivisible integral class $u \in H^1(M; \mathbf{Z})$ vanishes on $[\gamma_1], \ldots, [\gamma_k]$ and hence on all $[\gamma]$. We will deduce that ρ isn't transitive.

Let $\pi : \tilde{M} \to M$ be the connected $\mathbf{Z}$-cover of M determined by u. Then ρ_t lifts to a flow $\tilde{\rho}_t$ that commutes with $\mathbf{Z}$ and a periodic ρ-orbit lifts to a periodic $\tilde{\rho}$ orbit of the same period. We denote $\tilde{\rho}_t x$ by tx.

Let $\theta : M \to \mathbf{R}/\mathbf{Z}$ represents u, i.e., $\theta^*[dt] = u$. Then θ lifts to $\tilde{\theta} : \tilde{M} \to \mathbf{R}$. We show

LEMMA. $|\tilde{\theta}(tx) - \tilde{\theta}(x)|$ *is uniformly bounded over all* $x \in M$, $t \in \mathbf{R}$.

Proof. We may take x so that the orbit of x passes only through the interior of elements of $\mathcal{M}$. As every ρ orbit meets $\mathcal{M}$ in bounded time, we may assume x and tx lie over $\mathcal{M}$. For each t_i for which $t_i x \in \pi^{-1}\mathcal{M}$, $0 \le t_0 < t_1 < \cdots < t_n = t$, let $s(i) \in \mathcal{M}$ be the symbol associated to $\pi(t_i x)$.

Choose i_1 as large as possible with $s(i_1) = s(0)$, then i_2 is large as possible with $s(i_2) = s(i_1 + 1)$, etc., obtaining a sequence $0 \le i_1 < i_2 < \cdots \le i_m = n$. Let $i_0 = -1$. Then for $1 \le j \le m$, the sequence $s(i_{j-1} + 1), \ldots, s(i_j - 1), s(i_j)$ is a closed loop in $\mathcal{M}$ and determines a closed orbit passing through $y_j \in s(i_j)$ of some period p_j (not necessarily the minimum period of y_j). We lift y_j to the point $\tilde{y}_j \in \pi^{-1}(y_j)$ nearest to $a_j x$, where $a_j = t_{(i_{j-1})+1}$. Let $b_j = t_{i_j}$. Then

$$\tilde{\theta}(tx) - \tilde{\theta}(x) = \sum_{j=1}^{m} \tilde{\theta}(b_j x) - \tilde{\theta}(a_j x) + \sum_{j=2}^{m} \tilde{\theta}(a_j x) - \tilde{\theta}(b_{j-1} x).$$

Each term in the second sum is the variation of θ between consecutive intersections of a flowline with $\mathcal{M}$, hence bounded. The points $\pi(a_j x)$ and $\pi(y_i)$ have the same intersections with $\mathcal{M}$ for the first $i_j - i_{j-1}$ places and so stay close for that interval ($\mathcal{M}$ was chosen fine). Thus $b_j x$ is near y_j, and each term in the first sum is bounded. Finally $m \le \operatorname{card}(\mathcal{M})$, which proves the lemma.

To conclude the proof, we study the N-fold cover $M_N = \tilde{M}/N\mathbf{Z}$ of M and the induced flow $\psi_N = \tilde{\rho}/N\mathbf{Z}$ where N is chosen so that $N/3$ is a bound for the preceding lemma. Clearly M_N is compact and connected and ψ_N is Anosov (or pseudo-Anosov). Let $\theta_N : M_N \to \mathbf{R}/N\mathbf{Z}$ be the map induced by θ. As θ_N represents a nontrivial cohomology class, it is surjective. However, the values of θ_N over any trajectory of ψ_N are not dense (they occupy $\le 2/3$ of the image). Thus ψ_N isn't transitive, so $\Omega(\psi_N) \ne M_N$. But the density of closed orbits in M implies closed orbits are dense in M_N as well. This contradiction shows that u cannot exist, so $[\gamma]$ span as desired. Q.E.D.

6. Flow equivalence and Morse–Smale maps

We will show that flow-equivalence may be used to identify and strengthen an invariant of Asimov's for the suspension ρ of a Morse–Smale map. If $C_1, \ldots, C_r$

are the closed orbits of ρ, Asimov denotes the fixed point index of the Poincare map near C_i by $\epsilon_i = \pm 1$ and defines the geometric index of ρ to be $J(\rho) = \sum_{i=1}^r \epsilon_i [C_i] \in H$. In [A, proposition A] it is shown that if $f, g : K \to K$ are isotopic Morse-Smale maps that preserve stable and unstable orientations at their periodic points, their suspended flows ρ, ψ satisfy $J(\rho) = J(\psi)$.

In fact, $J(\psi)$ may be identified as follows.

THEOREM I. *For ψ as above, $J(\psi) = (-1)^k \chi(\psi^\perp)$, where $\chi(\psi^\perp)$ is the Euler class of the normal bundle $\psi^\perp$ and $k = \dim K$.*

Proof. If $u \in H^1(M; \mathbf{Z})$ is dual to K, then

$$u(J(\psi)) = \sum_{i=1}^{r} \epsilon_i u[C_i] = \sum_{i=1}^{r} \epsilon_i p_i$$

where p_i is the period of points $C_i \cap K$ for g. Also, $\epsilon_i = (-1)^{u_i}$, where u_i is the dimension of the unstable manifold of points in $C_i \cap K$. Thus $u(J(\psi)) = \sum_{\mathrm{Per}\,(g)} (-1)^{u(p)}$, where u is the unstable dimension of the periodic point p. By the Morse-Smale inequalities [*Sm*], $\sum_p (-1)^{k-u(p)} = \chi(K)$. This gives $u(J(\psi)) = (-1)^k \chi(K)$. The Euler class $\chi(\psi^1) \in H$ satisfies $u(\chi(\psi^\perp)) = \chi(K)$. Thus $u(J(\psi)) = u((-1)^k \chi(\psi^\perp))$.

Note that the condition $u(c_j) > 0$ is satisfied. Choosing a set of integral vectors $u_1, \ldots, u_\beta \in H^1(M; \mathbf{Z})$ that are close to u in angle but span $H^1(M; \mathbf{R})$, one obtains $u_i(\{c_j\}) > 0$, $i = 1, \ldots, \beta = \dim H^1(M; \mathbf{R})$. The arguments of the previous paragraph apply equally well to cross-sections K_i dual to u_i, and thus $u_i(J(\rho)) = u_i((-1)^k \chi(\rho^\perp))$ for $i = 1, \ldots, \beta$. As the u_i's span $H^1(M; \mathbf{R}) = H^* \otimes R$ and H is torsion free, the proposition follows. Q.E.D.

Consequently, Morse-Smale flows ρ, ψ with cross-section will have the same geometric index even if they are only assumed homotopic through nonsingular vector fields (for then $\chi(\rho^\perp) = \chi(\psi^\perp)$).

REFERENCES

[A] ASIMOV, D. *Flaccidity of geometric index for nonsingular vector fields*, Comm. Math. Helv. *52* (1977) 161–175.

[B1] BOWEN, RUFUS. *Symbolic dynamics for hyperbolic flows*. Amer. J. Math. *95* (1973), 429–460.

[B2] ——. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. Springer-Verlag LNM 470, 1975.

[B3] ——. *Topological entropy and Axiom A*, Global Analysis (Proc. Symp. Pure Math XIV, Berkeley, Calif. 1968) AMS, Providence 1970.

[DGS] DENKER, M., C. GRILLENBERGER and K. SIGMUND. *Ergodic theory on compact spaces*, Springer-Verlag LNM 527, 1976.

[FLP] FATHI, A., F. LAUDENBACH and V. POENARU (Eds.) *Travaux de Thurston sur les Surfaces – Seminaire Orsay*. Asterisques *66–67* (1979).
[Fr] FRANKS, JOHN. *Anosov diffeomorphisms*, Global Analysis (op. cit.).
[F1] FRIED, DAVID. *Cross-sections to flows*, Ph.D. thesis, U. Calif. Berkeley, 1976.
[F2] ——. *Geometry of cross-sections to flows*, to appear in Topology.
[Fu] FULLER, F. B. *The existence of periodic orbits*. Annals of Math. *57* (1953) 229–230.
[G] GALLOVOTTI, G. *Zeta functions and basic sets*, IHES preprint.
[H] HIRSCH, MORRIS. *Anosov maps, polycyclic groups and homology*, Topology *10* (1971) 177–183.
[M1] MANNING, A. *Anosov diffeomorphisms on nilmanifolds*. Proc. AMS *38* (1973) 423–426.
[M2] ——. *Axiom A diffeomorphisms have rational zeta functions*, Bull. London Math. Soc. *3*, (1971) 215–220.
[P] PLANTE, J. *Homology of closed orbits of Anosov flows*, Proc. AMS *37* (1973) 297–300.
[S] SCHWARTZMAN, S. *Asymptotic cycles*, Annals of Math. *66*, (1957) 270–284.
[Sm] SMALE, STEPHEN. *Differentiable dynamical systems*, Bull. AMS *73* (1967) 747–817.
[T] THURSTON, W. P. *On the geometry and dynamics of diffeomorphisms of surfaces*, preprint.
[W] WILLIAMS, R. F. *The structure of Lorenz attractors*, Publ. Math. IHES *50* (1980).

University of California
Santa Cruz, California 95064
USA

Received November 18, 1980/April 26, 1982