Séminaire de Théorie spectrale et géométrie

MARK POLLICOTT

Symbolic dynamics and geodesic flows

Séminaire de Théorie spectrale et géométrie, tome 10 (1991-1992), p. 109-129
<http://www.numdam.org/item?id=TSG_1991-1992__10__109_0>

Séminaire de théorie spectrale et géométrie
CHAMBÉRY-GRENOBLE
1991–1992 (109–129)

SYMBOLIC DYNAMICS AND GEODESIC FLOWS

by *Mark POLLICOTT*

Introduction

In these lectures we will try to give a coherent picture of the basic "symbolic dynamic" approach to geodesic flows (on compact manifolds with negative sectional curvatures).

The basic theory, as originally developed by Sinai, Bowen and Ruelle applies to more general differentiable flows (the so called "Anosov" and "Axiom A" flows) – but to avoid too much abstraction we prefer to keep geodesic flows as our basic context.

Some recent references are :

(a) *Ergodic theory, symbolic dynamics and hyperbolic spaces*, (ed. T. Bedford, M. Keane and C. Series), Proc. Conf. at ITCP (Trieste), April 1989.

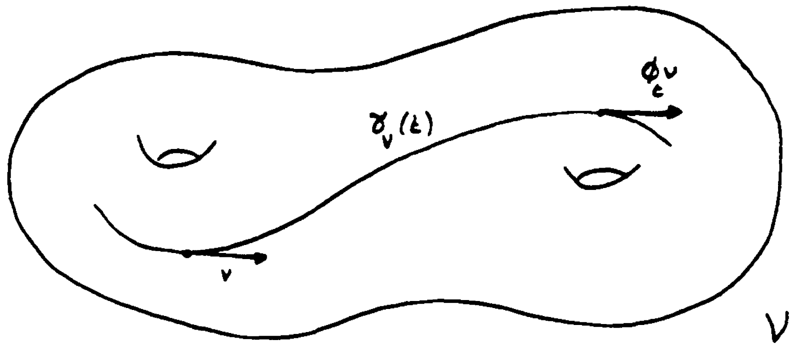
(b) W. Parry & M. Pollicott, *Zeta functions and closes orbit structure of hyperbolic flows*, Astérisque 186–187 (henceforth, denoted [PP] in these notes).

Basic setting. — We shall always let M denote a compact C^∞ Riemannian manifold (we shall always denote its dimension by $n = \dim M$). We shall let $\phi_t : M \longrightarrow M$ denote a C^∞ flow (*i.e.* one-parameter family of diffeomorphisms such that $\phi_t \circ \phi_s = \phi_{t+s}$, $t, s \in \mathbf{R}$).

Main examples. — Assume that V is a compact ℓ-dimensional C^∞ Riemannian manifold, with *all* sectional curvatures $\leqslant c < 0$. If $\|\ \|$ is the Riemannian metric for V, the unit tangent bundle (sphere bundle) is denoted by :

$$SV = \{V \in TM \mid \|V\| = 1\} .$$

We let $M = SV$ (which is $n = (2\ell - 1)$ dimensional) and define the *geodesic flow* $\phi_t : M \longrightarrow M$ by $\phi_t v = \dot\gamma_v(t)$, where $\gamma_v : \mathbf{R} \longrightarrow V$ is the geodesic with $\dot\gamma_v(0) = V$.



A. Sections and the Markov property

A basic tool in dynamical systems is to reduce a (continuous) flow to a (discrete) transformation by introducing a finite number of "Poincaré sections" transverse to the direction of the flow $\phi_t : M \to M$.

Assume that we have disjoint $(n-1)$-dimensional (closed) sections $T_1, \ldots, T_k \subseteq M$ (usually assumed to be contained in C^∞ discs $T_i \subset D_i$ transverse to the flow) with the following properties :

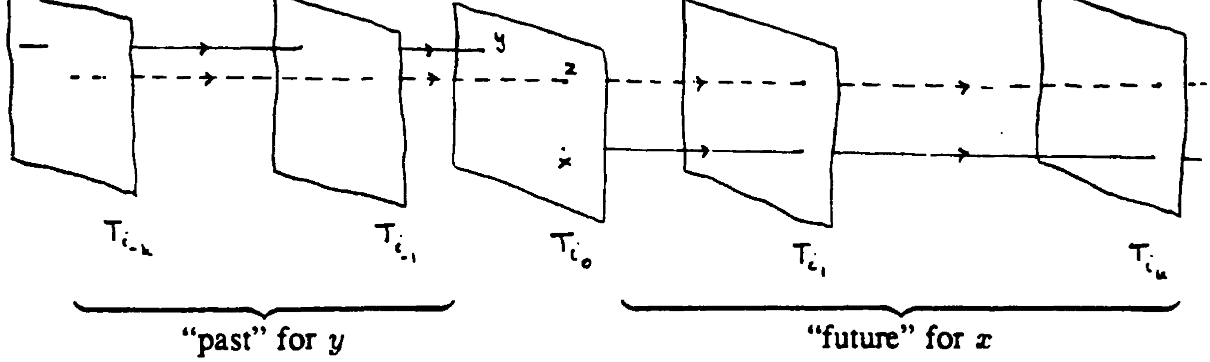
Simple properties.

(i) $T_i = \mathrm{cl}(\mathrm{int}\, T_i)$ ("Proper")
(where the closure and interior are in the topology of the $(n-1)$ dimensional discs).

(ii)$_\varepsilon$ $M = \bigcup_i \phi_{[0,\varepsilon]} T_i$ ("Sections sweep out M")
(where $\varepsilon > 0$ is some "small" number whose role will become apparent in the construction).

Special property.

(iii) If the *future* orbit $(\phi_t x,\ t>0)$ of $x \in \mathrm{int}\, T_{i_0}$ passes through $\mathrm{int}\, T_{i_k}$, $k = 1, 2, 3, \ldots$ (in sequence) and the *past* orbit $(\phi_t y,\ t<0)$of $y \in \mathrm{int}\, T_{i_0}$ passes through $\mathrm{int}\, T_{i_{-k}}$, $k = 1, 2, 3, \ldots$ (in sequence) then there exists a single point $z = z(x, y) \in \mathrm{int}\, T_{i_0}$ which has both properties ("Markov property").



DEFINITION. — We call a family $\mathcal{T} = \{T_1, \ldots, T_k\}$ of Poincaré sections satisfying (i), (ii)$_\varepsilon$, (iii) a *family of Markov sections* (If also we have $\max_{i \leqslant i \leqslant k} \operatorname{diam}(T_i) \leqslant \varepsilon$ then we say it is of "size ε").

Note. — There is an equivalent (and less intuitive) version of property (iii) which is commonly used. We shall come back to this below.

The main existence theorem we need is the following

THEOREM (Bowen-Ratner). — *If $\phi_t : M \longrightarrow M$ is a geodesic flow (for a compact C^∞ manifold with negative sectional curvatures) then there exists a family of Markov sections (of arbitrarily small size).*

Warning. — In general, there is nothing canonical/unique about families of Markov sections (This will be apparent from the sketch of the proof below). The exception is the fundamentally different construction of Adler-Flatto/Series for surfaces with $\kappa = -1$ (*cf.* Adler-Flatto, B.A.M.S., vol. 25 (1991), 229–334).

B. The proof of the Bowen-Ratner theorem

We want to give a sketch of the construction of a family $\mathcal{T}$ (as promised by the Bowen-Ratner theorem). The reason that negative sectional curvature is important is that the flow $\phi_t : M \longrightarrow M$ has a special "hyperbolic" structure, which we shall describe below :

A *stable manifold* (or horocycle), through $x, \in M$ is the set

$$W^s(x) = \left\{ y \in M \mid d(\phi_t x, \phi_t y) \longrightarrow 0 \text{ as } t \longrightarrow +\infty \right\} ,$$

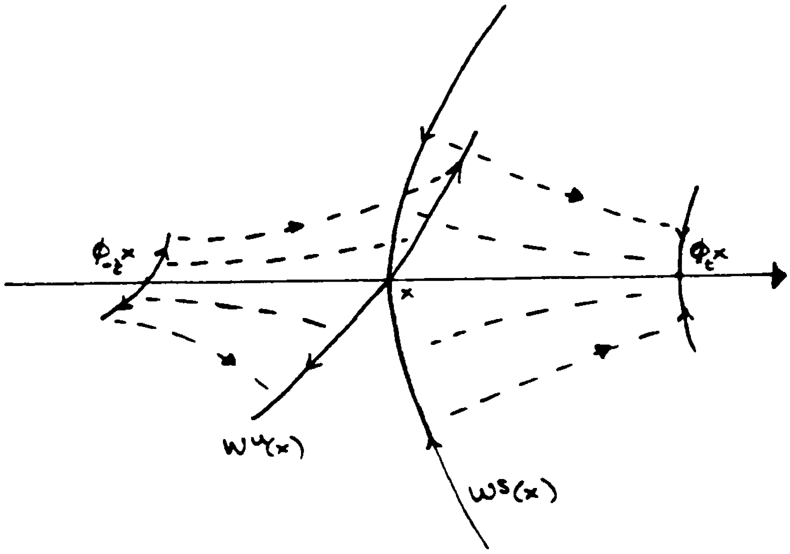
and the *unstable manifold* (or horocycle), through $x \in M$, is the set

$$W^u(x) = \left\{ y \in M \mid d(\phi_{-t} x, \phi_{-t} y) \longrightarrow 0 \text{ as } t \longrightarrow +\infty \right\} .$$

Note. — For geodesic flows these are C^∞ immersed submanifolds and $W^s = \{W^s(x) \mid x \in M\}$, $W^u = \{W^u(x) \mid x \in M\}$ are continuous foliations.

(Observe that $\phi_t W^s(x) = W^s(\phi_t x)$, $\phi_t W^u(x) = W^u(\phi_t x)$, $(\forall t \in \mathbf{R})$).

DEFINITION. — We call the flow $\phi_t : M \longrightarrow M$ *hyperbolic* if the convergence in $W^s(x)$, $W^u(x)$ is exponentially fast (*i.e.* $\exists C$, $\lambda > 0$ s.t. $\|D\phi_t \mid T_x W^s(x)\| \leqslant Ce^{-\lambda t}$, $\|D\phi_{-t} \mid T_x W^u(x)\| \leqslant C^{-\lambda t}$, $t \geqslant 0$),



i.e. the flow "contracts lengths" exponentially fast on W^s, and "expands lengths" exponentially fast on W^u.

Of course, the main examples are geodesic flows :

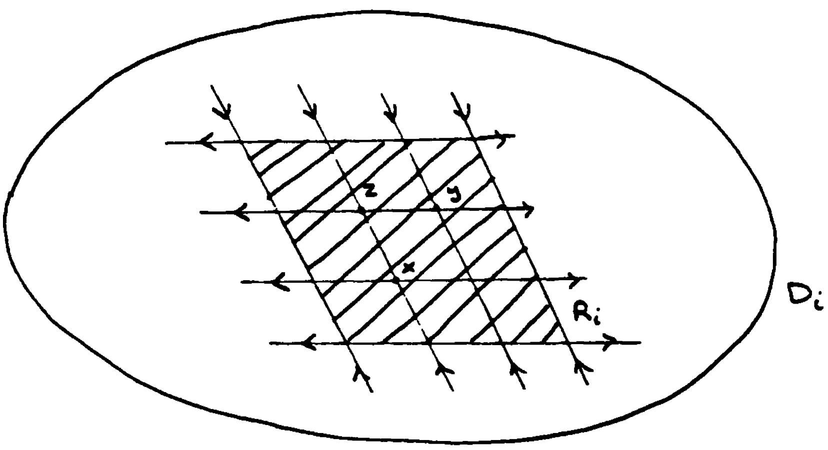
PROPOSITION (Anosov). — *Geodesic flows $\phi_t : M \to M$ (for compact C^∞ manifolds with negative sectional curvatures) are hyperbolic.*

(The proof is based on the study of stable and unstable Jacobi fields, *cf.* Anosov-Sinai, Russian Math. Surveys, vol. 22(5) (1967), 103–167).

Armed with these definitions, we can proceed to the proof of the Bowen-Ratner theorem :

Sketch proof of theorem. — (This proof also works for general hyperbolic flows).

Step 1. Choose disjoint C^∞ discs $D_1, \ldots, D_N$ transverse to the flow. Choose "Rectangles" $T_i^0 \subseteq D_i$ (*i.e.* closed sets such that $\forall x, y \in R_i$ the intersection $z = [x, y] = \mathrm{Proj}_{D_i} W^s(x) \cap \mathrm{Proj}_{D_i} W^u(y) \in R_i$)

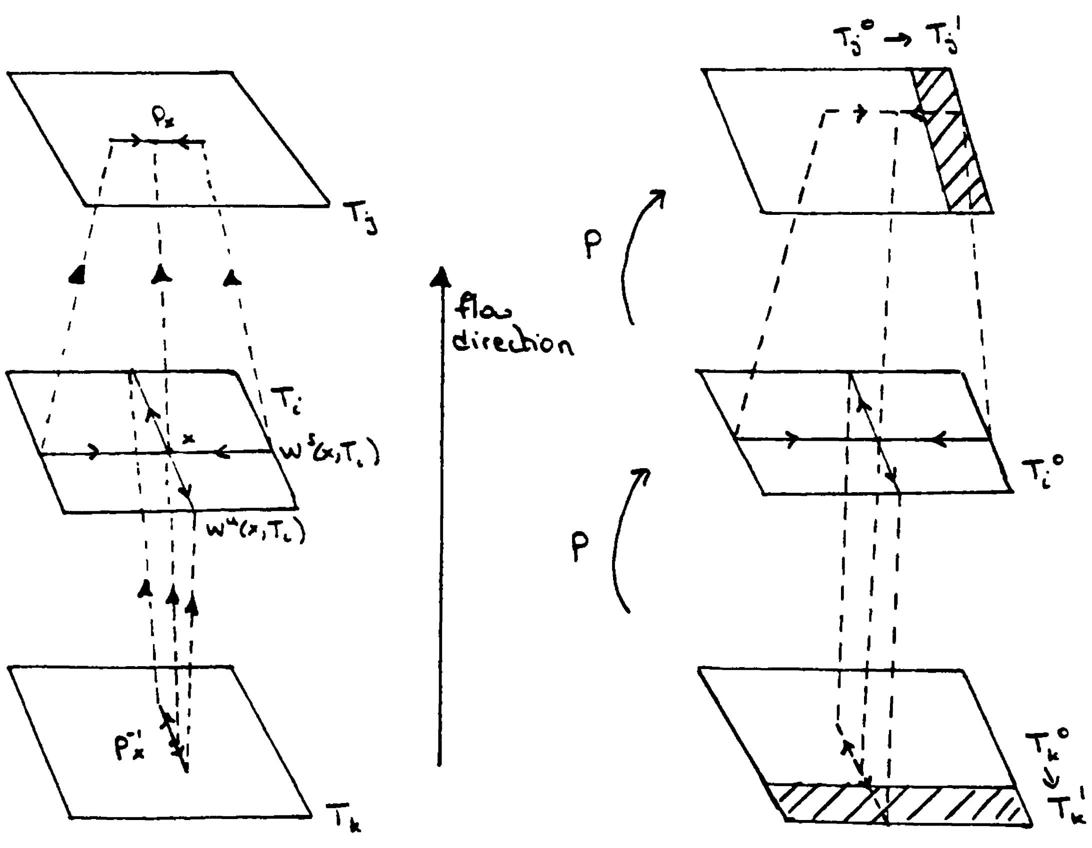


Assume that $\{T_i^0 \mid i = 1, \dots, N\}$ satisfy (i) and (ii)$_\varepsilon$. These are our "zeroth" approximation. They are improved by an inductive argument.

Step 2. For a family $\mathcal{T} = \{T_1, \dots, T_k\}$ to satisfy (iii) it *suffices* that the Poincaré first return map $P : \bigcup_{i=1}^{k} \operatorname{int} T_i \longrightarrow \bigcup_{i=1}^{k} \operatorname{int} T_i$ satisfy :

$$\text{(iii)}' \qquad \begin{cases} x \in \operatorname{int} T_i,\ Px \in \operatorname{int} T_j \Longrightarrow W^s(x, T_i) \subseteq W^s(P_x, T_j) \\ x \in \operatorname{int} T_i,\ P^{-1}x \in \operatorname{int} T_k \Longrightarrow P^{-1} W^u(x, T_i) \subseteq W^u(x, T_k) \end{cases}$$

(where $W^s(x, T_i) = [x, T_i]$, $W^u(x, T_i) = [T_i, x]$)



We enlarge rectangles $\{T_1^0, \ldots, T_N^0\}$ as illustrated, so that *some* return time has this property "to first order" (as in diagram). Iterate this procedure to get rectangles $\{T_1^\ell, \ldots, T_N^\ell\}$, $\ell \geqslant 0$. By hyperbolicity, $\{T_i^\infty := \bigcup_{\ell=0}^{\infty} T_i^\ell\}$ exist, and property (iii)′ holds for some return time.

Step 3. By replacing $\{T_i^\infty\}$ by $\{T_i\} := \{\mathrm{cl}(\mathrm{int}\, T_{i_0} \cap P^{-1}\, \mathrm{int}\, T_{i_1} \cap \cdots \cap P^{-n}\, \mathrm{int}\, T_{i_m})\}$ (for sufficiently large n) we can assume (iii)′ holds for the *first* return map. ∎

Reference. — Detailed proofs occur in Bowen, Amer. J. Math., vol. 75, (1973) and Ratner, Israel J. Math. (1973) and a sketch appears in Appendix III of Parry and Pollicott [PP]. A good general reference is Alexseev-Jakobson, Physics reports, vol. 75.

C. Modelling the Poincaré map

Given sections $\mathcal{T} = \{T_1, \ldots, T_k\}$, we want to introduce a "symbolic" model for the Poincaré map $P : \bigcup_{i=1}^{k} T_i \longrightarrow \bigcup_{i=1}^{k} T_i$. The *symbols* are the indices $\{1, \ldots, k\}$ for the sections.

Given $z \in \mathrm{int}\, T_{x_0}$, assume that $P^n z \in \mathrm{int}\, T_{x_n}$, $x_n \in \{1, \ldots, k\}$, $\forall n \in \mathbf{Z}$. We associate to z the sequence $(x_n)_{n=-\infty}^{+\infty}$.

We define a $k \times k$ matrix A with entries

$$A(i,j) = \begin{cases} 1 & \text{if } \; P(\mathrm{int}\, T_i) \cap \mathrm{int}\, T_j = \emptyset \\ 0 & \text{if } \; P(\mathrm{int}\, T_i) \cap \mathrm{int}\, T_j \neq \emptyset \end{cases}$$

$(1 \leqslant i, j \leqslant k)$ and a *space* of sequences

$$X_A = \left\{ x = (X_n)_{n=-\infty}^{+\infty} \mid A(x_i, x_{i+1}) = 1,\ i \in \mathbf{Z} \right\}$$

we can give this a family of *metrics* d_θ $(0<\theta<1)$ defined by $d_\theta(x,y) = \theta^n$, where $n = n(x,y) = \sup\{m \geqslant 0 \mid x_i = y_i,\ |i| \leqslant m\}$ (and $d_\theta(x,x) \equiv 0$). We can define a *homeomorphism* $\sigma : X_A \longrightarrow X_A$ by shifting sequences one place to left *i.e.* $\sigma((x_n)_{n=-\infty}^{+\infty}) = (x_{n+1})_{n=-\infty}^{+\infty}$.

Note. — With this metric, two sequences $x = (x_n)_{n=-\infty}^{+\infty}$, $y = (y_n)_{n=-\infty}^{+\infty} \in X_A$ are "close" if they agree in a larger number of terms.

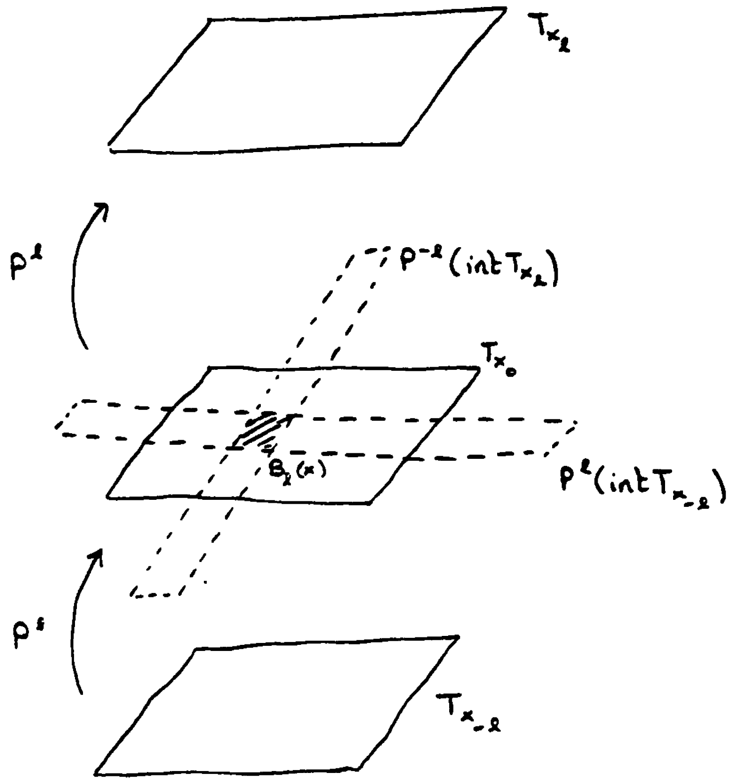
The following properties are easily checked :

LEMMA.

(a) (X_A, d_θ) is compact ;

(b) The sets $U(x_{-n}, \ldots, x_n) = \{y \in X_A \mid y_i = x_i,\ -n \leqslant i \leqslant n\}$ (which are both open and *closed) are a sub-basis for (X_A, d_θ).*

Finally, we define a map $\pi : X_A \longrightarrow \bigcup_{i=1}^{k} T_i$ by $\{\pi(x)\} = \bigcap_{n=-\infty}^{+\infty} \overline{P^{-n}(\operatorname{int} T_{x_n})}$. To see this is well-defined, let $B_\ell(x) = \bigcap_{n=-\ell}^{\ell} \overline{P^{-n}(\operatorname{int} T_{x_n})}$, $\ell \geqslant 0$ and observe $\bigcap_{\ell=0}^{\infty} B_\ell(x) = \bigcap_{n=-\infty}^{+\infty} \overline{P^{-n}(\operatorname{int} T_{x_n})}$. Using (iii) we can check $B_\ell(x) \neq \emptyset$ ($\ell \geqslant 0$), and since $B_1(x) \supseteq B_2(x) \supseteq B_3(x) \supseteq \cdots$, and each is compact, we see that the intersection is *non-empty*.



To see that the intersection is a *single* point, we use the hyperbolicity to see that $\operatorname{diam}(B_\ell) \to 0$ (in fact, $\operatorname{diam}(B_\ell) \leqslant \operatorname{const} \gamma^n$, for some $0 < \gamma < 1$ $(*)$).

The *shift* $\sigma : X_A \to X_A$ is a good model for $P : \bigcup_{i=1}^{k} T_i \longrightarrow \bigcup_{i=1}^{k} T_i$, as shown by the following :

LEMMA.

(i) π is Hölder continuous (i.e. $\exists C, \alpha > 0$ such that $d(\pi x, \pi y) \leqslant C[d_\theta(x,y)]^\alpha$) ;

(ii) π is "usually" injective – in fact, one-one on a dense Baire set (i.e. countable intersection of open dense sets) and surjective ;

(iii) π is always bounded-one (i.e. $\exists N \geqslant 1$ such that $\operatorname{card}\{\pi^{-1}(z)\} \leqslant N$, *$z \in \bigcup_{i=1}^{k} T_i$, and we can even take $N = k^2$) ;*

(iv) $P\pi = \pi\sigma$ *if* $\pi(x)$, $\pi(\sigma x) \in \bigcup_{i=1}^{k} \operatorname{int} T_i$.

Remarks.

(a) Notice that the choice of θ is unimportant in (i) – if we replace θ by θ' then we just replace α by $\alpha' = \frac{\alpha \log \theta}{\log \theta'}$;

(b) π is "usually" injective in a measurable sense *i.e.* it fails to be injective on a subset of $\bigcup_{i=1}^{k} T_i$ of zero Lebesgue (*i.e.* volume) measure.

Sketch proof of lemma. — The Holder continuity in (i) comes from the hyperbolicity property $(*)$. For part (ii), let $B = \{z \in \bigcup_{i=1}^{k} T_i \mid P^n z \in \bigcup_{i=1}^{k} \operatorname{int} T_i, \forall n \in \mathbf{Z}\}$ then B is a Baire set, and for $z \in B$ we construct a unique element $x = (x_n)_{n=-\infty}^{+\infty}$ by $P^n z \in \operatorname{int} T_{x_n}$, then $x = \pi^{-1} z$. Furthermore, by construction $\pi(X_A) \supseteq B$ and since : X_A is compact ; π is continuous (thus $\pi(X_A)$ is closed) ; and B is dense we see that $\pi(X_A) = \bigcup_{i=1}^{k} T_i$ (*i.e.* π is surjective) (Part (iii) is easy, but has a trick – so we shall skip it). Finally, (iv) is obvious from the definition of π.

Note. — The subshift $\sigma : X_A \to X_A$ gives only the basic configuration of the sections. Next we must introduce a function which "encodes" information about lengths.

D. Modelling the geodesic flow

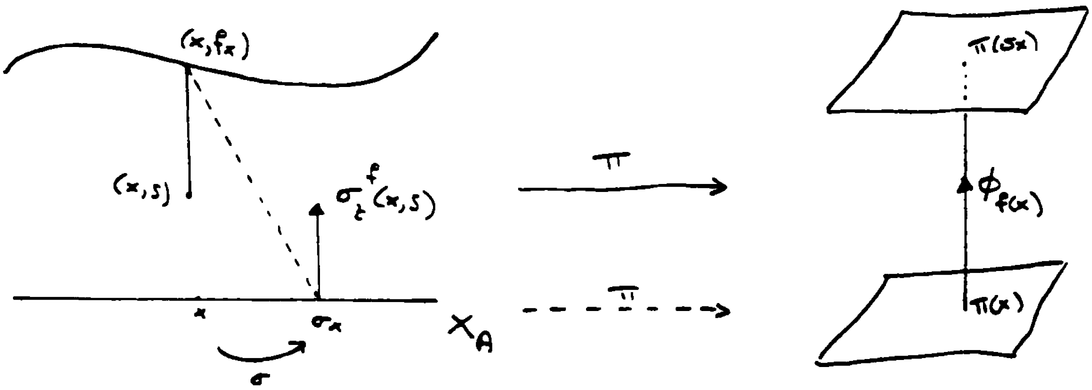
It is now a simple matter now to build a model for the whole flow (from the model for the Poincaré map).

The symbolic model consists of a metric space :

$$X_A^f = \left\{(x,t) \in X_A \times \mathbf{R}^+ \mid 0 \leqslant t \leqslant f(x)\right\} / \sim$$

for a Hölder continuous *roof function* $f : X_A \to \mathbf{R}^+$ (*i.e.* $\exists C, \alpha > 0$ such that $|f(x) - f(y)| \leqslant C[d_\theta(x,y)]^\alpha$) *where* we identify the "top" and "bottom" by $(x, r(x)) \sim (\sigma x, 0)$.

We define a *suspended flow* $\sigma_t^f : X_A^f \to X_A^f$, by $\sigma_t^f(x,s) = (x, s+t)$ (if $0 \leqslant s$, $s+t \leqslant f(x)$, and using the identification for other values of t).



We choose $f : X_A \to \mathbf{R}^+$ by $\phi_{f(x)}\pi(x) = \pi(\sigma x)$ (**). We can extend $\pi : X_A^f \to \bigcup_{i=1}^{k} T_i$ to a map $\pi : X_A^f \to M$, by $\pi(x,t) = \phi_t \pi(x)$ (where (**) takes care of the identification $\sim$). Again, we see σ_t^f is a "good" model for ϕ_t in the following sense.

LEMMA.

(i) π *is Hölder continuous ;*

(ii) π *is one-one on a dense Baire set, and surjective ;*

(iii) π *is always bounded-one ;*

(iv) $\phi_t \pi = \pi \sigma_t^f$, $t \in \mathbf{R}$.

E. Simpler roof functions

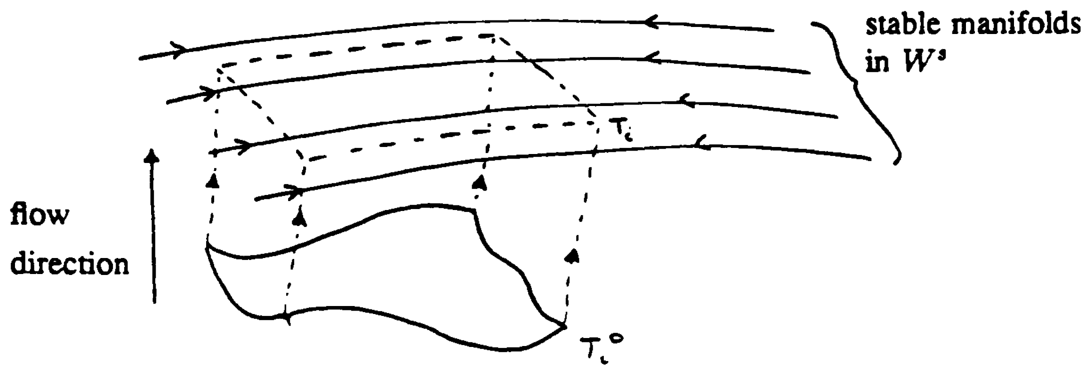
Since we have so much flexibility in choosing our Poincaré sections, we want to make a choice that will simplify life later.

DEFINITION. — *We call* $g \in C^0(X_A, \mathbf{R})$ a function of the future *if* $g(x) = g(y)$ *whenever* $x_i = y_i$ *for* $i \geqslant 0$ *(i.e.* $g(x) = g(x_0, x_1, x_2, \dots)$ *depends* only *on* x_i, $i \geqslant 0$*).*

Notational comment. — The terms x_i, $i \geqslant 0$, are the "future" (and present) and the terms x_i, $i < 0$ are the "past". The map $\sigma : X_A \to X_A$ moves the "future" to the past (like time !).

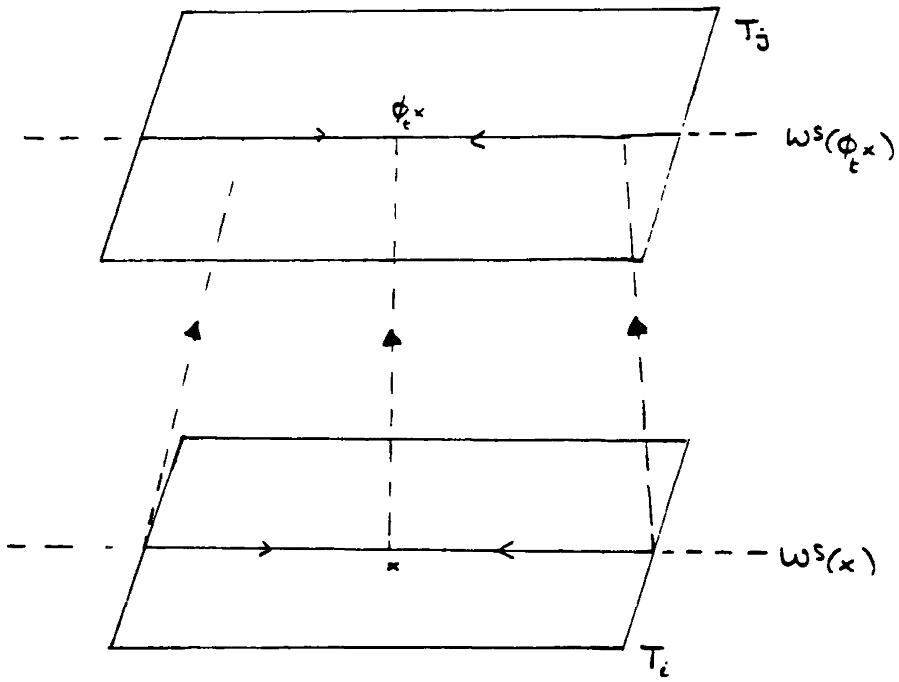
LEMMA. — *For a suitable choice of sections* $\mathcal{T} = \{T_1, \dots, T_k\}$ *we can assume* $f : X_A \to \mathbf{R}^+$ *is a Hölder continuous function of the future.*

Sketch proof. — Starting from a family of sections $\mathcal{T}^0 = \{T_1^0, \dots, T_k^0\}$ (satisfying (i), (ii) and (iii)) we "shear" them in the flow direction so that each section T_i^0 is "foliated" by stable manifolds from W^s. The resulting section is denoted T_i



Notice that :

(i) The Poincaré return time is constant *along* stable manifolds (using the property : $\phi_t W^s(x) = W^s(\phi_t x)$, $t \in \mathbf{R}$)



At the symbolic level, f is a "function of the future"

(ii) The new sections T *still* satisfy the important properties. However, we only expect each T_i to we as regular as the foliations W^s.

Fact (Anosov). — In general, the foliation W^s are Hölder continuous.

(For geodesic flows on surfaces, or with pinching conditions on sectional curvature, we get that W^s is C^1 (Hirsch-Pugh, J. Diff. Geom. (1975))).

Thus, we conclude each T_i is at *least* Hölder and this means $f : X_A \rightarrow \mathbf{R}^+$ is Hölder.

Remarks.

(i) In special cases like manifolds with *constant* negative sectional curvatures, the foliation W^s is C^ω and it is better to retain more geometry (*cf.* Ruelle, Invent. Math., vol. 34).

(ii) In some special cases (like co-compact manifolds) all the "interesting" behaviour is on a compact (Recurrent) subset $\Lambda \subseteq M$ of the flow. With some modifications, we get an analogous symbolic model.

F. Some analysis

The motivation for introducing a symbolic model for the flow is that there are some useful "tools" associated to shifts. However, there is first a preliminary "reduction" :

We let $X_A^+ = \{x = (x_n)_{n=0}^{\infty} \mid A(x_i, x_{i+1}) = 1,\ i \geqslant 0\}$ (*i.e.* we consider only the "future" part of sequences) and define a one-sided shift $\sigma^+ : X_A^+ \longrightarrow X_A^+$ by $\sigma^+((x_n)_{n=0}^{\infty}) = (x_{n+1})_{n=0}^{\infty}$ (*i.e.* shifting sequences one place to the left, forgeting the first term). We define metric(s) on X_A^+ by $d(x,y) = \theta^m$, where $x = (x_n)_{n=0}^{\infty}$, $y = (y_n)_{n=0}^{\infty}$ and $m = m(x,y) = \sup\{\ell \geqslant 0 \mid x_i = y_i,\ 0 \leqslant i \leqslant \ell\}$.

Note. — As with X_A, the space X_A^+ is compact and has a sub-basis of open-closed sets. σ^+ is still continuous (but not invertable!). There is an obvious map $\rho : X_A \longrightarrow X_A^+$ $(\rho((x_n)_{n=-\infty}^{+\infty}) = (x_n)_{n=0}^{\infty})$ and we see that we can identify "functions of the future" $g \in C^0(X_A, \mathbf{R})$ with functions $\bar{g} \in C^0(X_A^+, \mathbf{R})$ (*i.e.* $g = \bar{g} \circ \rho$).

We denote by $F_\theta := \{g : X_A^+ \longrightarrow \mathbf{R} \mid \exists C > 0,\ |g(x) - g(y)| \leqslant C \cdot d_\theta(x,y)\}$ (where we choose $0 < \theta < 1$ sufficiently small that $f \in F_\theta$).

LEMMA. — *F_θ is a Banach space with norm $||g||_\theta = ||g||_\infty + C(g)$ (where $C(g) = \sup_{x \neq y} \{\dfrac{|g(x) - g(y)|}{d(x,y)}\}$ is the "best" bound $C > 0$).*

(The proof is elementary).

At the very heart of our analysis is a bounded linear operator on this Banach space.

DEFINITION. — Given $\omega \in F_\theta$ we define the associated *Ruelle operator*

$$L_\omega : F_\theta \longrightarrow F_\theta \text{ by } (L_\omega g)(x) = \sum_{y:\sigma^+ y = x} e^{\omega(y)} \cdot g(y)$$

(where the summation is over at most k pre-images $(\sigma^+)^{-1}x = \{y = (i, x_0, x_1, \ldots) \mid A(i, x_0) = 1\}$).

Note. — If we assume $0 < \theta < 1$ and $\omega \in F_\theta$, then an easy exercise gives :

$$C(L_\omega g) \leqslant \text{const}\,||g||_\infty + \theta \cdot C(g) \tag{$*$}$$

where const $= \frac{C(e^\omega)}{1-\theta}$. Thus $L_\omega(F_\theta) \subseteq F_\theta$.

If we assume ω is complex valued then we replace F_θ by complex-valued functions (*i.e.* the "complexification" of F_θ).

The basic result on this operator is the following ;

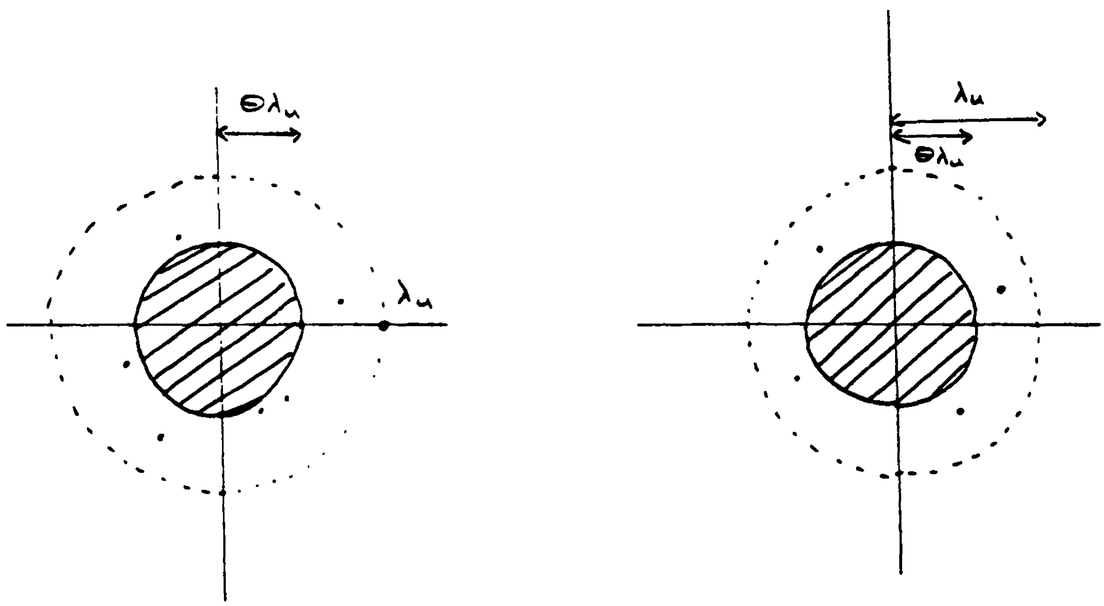
THEOREM (Ruelle operator theorem).

(i) If u real valued then L_u has a maximal positive eigenvalue $\lambda_u > 0$ (with a simple eigenfunction $h_u \geqslant 0$).

(ii) If $\omega = u + iv$ then the spectrum of L_ω consists of

(a) Isolated eigenvalues (with finite dimensional generalized eigenspaces) in the annulus $\lambda_u \geqslant |z| > \theta\lambda_u$;

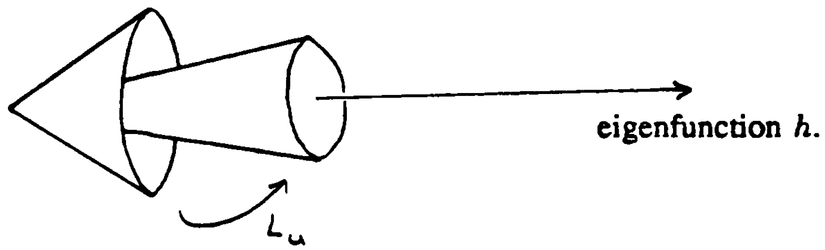
(b) The disc $\{z \mid |z| \leqslant \theta\lambda_u\}$.



$\mathcal{L}_u$ *spectrum* $\mathcal{L}_{u+iv}$ *spectrum*

Sketch proof.

(i) Showing the existence of a positive eigenvalue involves studying a positive cone $\mathcal{P}$ with $L_u : \mathcal{P} \subset \mathcal{P}$.



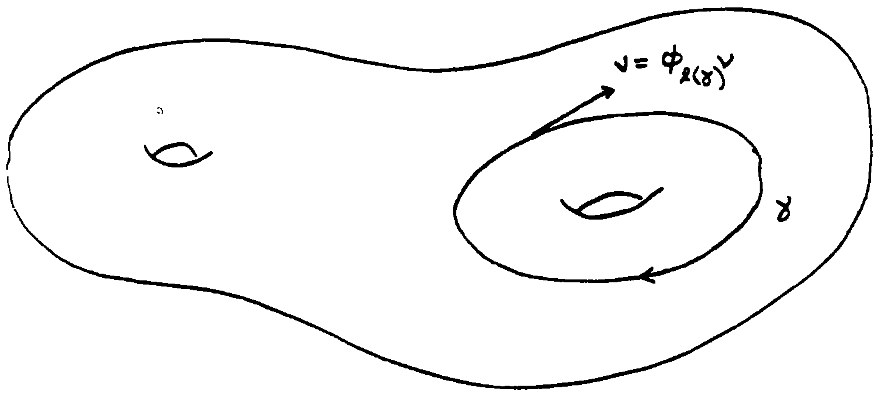
(ii) This follows from a version of the (Essential) spectral radius theorem, due to Nussbaum, *cf.* [PP].

DEFINITION. — We define $P(u) = \log \lambda_u$ to be the *Pressure function.*

Remark. — For practical applications, we let $\omega = -sf$, $s \in \mathbf{C}$.

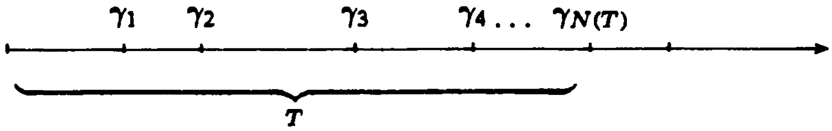
G. Topological entropy for the flows

The geodesic flow $\phi_t : M \to M$ has a countable number of closed orbits γ (of least period $\ell(\gamma)$). (These correspond precisely to the closed geodesics on the manifold).



Given $T > 0$, denote by $N(T) = \#\{\gamma \mid \ell(\gamma) \leqslant T\}$.

DEFINITION. — The *topological entropy* $h(\phi)$ of the flow $\phi_t : M \to M$ is defined by $h(\phi) = \overline{\lim_{T \to \infty}} \frac{1}{T} \log N(T)$.



Remark. — Thus $h(\phi)$ is a "growth rate" of closed orbits. There are many alternative definitions (e.g. "growth rate" of volume in the universal cover). More general definitions exist of the topological entropy of any continuous map on a compact space.

Similarly, we can define $\widetilde{N}(T)$ to be the number of closed orbits $\widetilde{\gamma}$ for $\sigma_t^f = X_A^f \to X_A^f$ with least period $\ell(\widetilde{\gamma}) \leqslant T$, and define its topological entropy by

$$h(\sigma^f) = \overline{\lim_{T \to \infty}} \frac{1}{T} \log \widetilde{N}(T) .$$

PROPOSITION (comparison of entropies). — $h(\phi) = h(\sigma^f)$.

Idea of proof. — It suffices to show that $\frac{\widetilde{N}(T)}{N(T)} \longrightarrow 1$ as $t \to \infty$. In fact, the "almost one-one" correspondence guarantees this.

Remark. — There are various alternative proofs.

We want to find a link between the Pressure (defined at the level of the shift $\sigma^+ : X_A^+ \to X_A^+$) and the topological entropy (for the suspended flow $\sigma_t^f : X_A^f \to X_A^f$), which allows us to apply our results on the Ruelle operator. The next section introduces the machinary for this.

H. Measures and the variational principle

We now move into the realm of "Ergodic Theory" where, in particular, all measures are probabilities and invariant under the appropriate transformation or flow.

DEFINITIONS. — A σ-invariant probability measure μ on X_A satisfies :

(i) $\mu(X_A) = 1$ (probability) ;

(ii) $\int F \, d\mu = \int F \circ \sigma \, d\mu$, $\forall F \in C^0(X_A, \mathbf{R})$ (invariance).
(and similarly for $\sigma^+ : X_A^+ \to X_A^+$).

A σ^f-invariant probability measure m on X_A^f satisfies :

(i) $m(X_A^f) = 1$;

(ii) $\int F \, dm = \int F \circ \sigma_t^f \, dm$, $(\forall t \in \mathbf{R})$, $\forall F \in C^0(X_A^f, \mathbf{R})$.

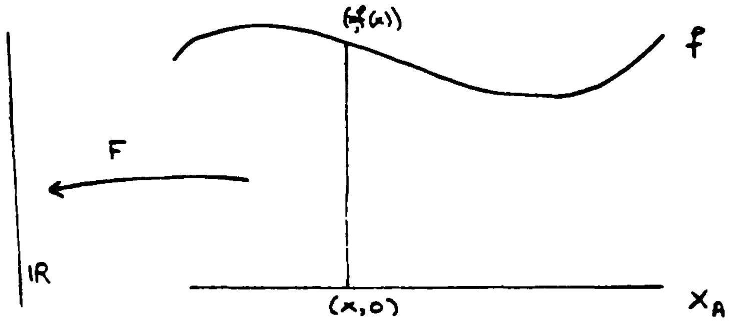
SIMPLE LEMMA.

(a) There is a bijection between invariant probability measures on X_A and X_A^+.

(b) There is a bijection between invariant probability measures on X_A and

X_A^f, *where*

$$\int_{X_A^f} F\, dm = \int_{X_A} \left\{ \int_0^{f(x)} F(x,t)\, dt \right\} d\mu(x) \Big/ \int f\, d\mu \;. \quad (\textit{written } dm = \frac{d\mu \times dt}{\int f\, d\mu}) \;.$$



The basic numerical "invariant" associated to invariant probability measures is "measure theoretic entropy".

In the case of a σ-invariant propability measure μ on X_A, we define the *measure theoretic entropy* $h(\sigma, \mu)$ by

$$h(\sigma,\mu) = \overline{\lim_{N\to\infty}} \frac{1}{N} \sum \mu[x_0, \dots, x_{N-1}] \log \mu[x_0, \dots, x_{N-1}] \geqslant 0 \qquad (*)$$

where the sum is over the sets

$$[x_0, \dots, x_{N-1}] = \left\{ y \in X_A \mid y_i = x_i,\ 0 \leqslant i \leqslant N-1 \right\} .$$

Remark. — Using a subadditivity trick, we can see that the limit in $(*)$ is an infimum (of continuous functions in μ).

The "heuristic interpretation" of entropy is that the larger the entropy, the more complicated (or interesting) the invariant measure.

The connection with the Ruelle operator is given by the following fundamental result :

THEOREM (variational principle). — *For any real valued $u \in F_\theta$*

$$P(u) = \sup \left\{ h(\mu,\sigma) + \int u\, d\mu \mid \mu = \sigma\text{-invariant probability measure} \right\} \qquad (*)$$

and there exists a unique measure μ (called the "equilibrium state") such that $P(u) = h(\mu,\sigma) + \int u\, d\mu$.

Note. — The measure μ comes from the eigenprojection $\mathcal{L}_u^* \nu = \lambda \nu$ by $d\mu/d\nu = h$ (where $\mathcal{L}_u h = \lambda h$).

Application. — The variational principle leads to a very useful result, which will shall refer back to later :

COROLLARY (Derivatives of Pressures).

(i) $\frac{d}{dt}P(u+tu')|_{t=0} = \int u'\, d\mu \quad (\mu = \text{equilibrium state for } u)$;

(ii) $\frac{d^2}{dt^2}P(u+tu')|_{t=0} = \lim_{n\to\infty} \frac{1}{n} \int (\sum_{i=0}^{n-1} u' \circ \sigma_i)^2\, d\mu(x)$.

Proof of corollary. — By perturbation theory, $t \mapsto L_{u+tu'} \mapsto \lambda_{u+tu'} \mapsto \log \lambda_{u+tu'} = P(u+tu')$ is analytic. The first derivative how comes from the variational principle :

$$\begin{cases} P(u+tu') \geqslant h(\mu,\sigma) + \int (u+tu')\, d\mu = P(u) + t \int u'\, d\mu \\ P(u-tu') \geqslant h(\mu,\sigma) + \int (u-tu')\, d\mu = P(u) - t \int u'\, d\mu \end{cases}$$

$$\Longrightarrow \frac{P(u+tu') - P(u)}{t} \geqslant \int u'\, d\mu \geqslant \frac{P(u) - P(u-tu')}{t} \quad (\text{let } t \to 0) .$$

(The second derivative is also easy, *cf.* [PP]).

I. A symbolic characterisation of entropy

We begin with the following "classical" result :

PROPOSITION.

(i) The σ^f-invariant measure $dm = d\mu \times dt / \int f\, d\mu$ has measure theoretic entropy $h(m) = h(\sigma, \mu) / \operatorname{int} f \cdot d\mu$ (Abramov) ;

(ii) The topological entropy is given by

$$h(\sigma^f) = \sup \{ h(m) \mid m = \sigma^f - \textit{invariant} \} \quad (\text{Adler}) .$$

This brings us to the following important characterisation of topological entropy :

THEOREM (Symbolic characterisation of $h(\sigma^f)$). — *The topological entropy $h(\sigma^f)$ is the unique value $t = h(\sigma^f)$ such that $P(-tf) = 0$.*

Proof of Theorem. — We use the variational principle to write :

$$P(-tf) = 0 \geqslant h(\sigma,\mu) - t \int f\, d\mu, \ \forall \sigma\text{-invariant } \mu ,$$

$$\Longleftrightarrow t \geqslant \frac{h(\sigma,\mu)}{\int f\, d\mu} = h(m), \ \forall \sigma^f\text{-invariant } dm = \frac{d\mu \times dt}{\int f\, d\mu} \ (\text{Abramov})$$

(with equality when μ = equilibrium state for $-tf$)

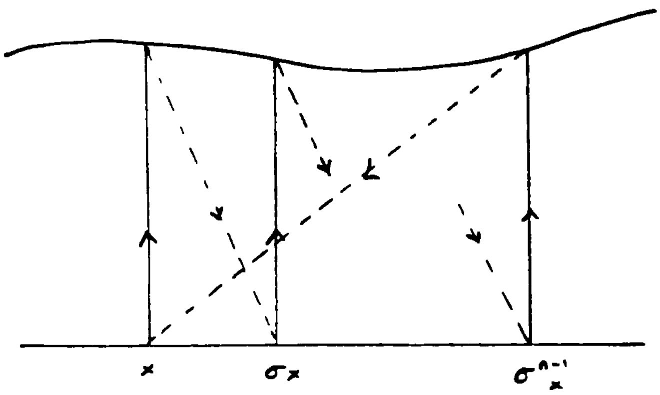
$$\Longleftrightarrow t = h(\sigma^f) \quad (\text{Adler}) .$$

Comment. — The analyticity of the map $t \mapsto P(-tf)$ and this characterisation of h is central to smoothness results on entropy.

J. Periodic points and zeta functions

Our definition of topological entropy is in terms of the asymptotic behaviour of closed orbits and $P(t)$. We want to use the symbolic model (and, in particular, the spectrum of the Ruelle operators L_{-sf}, $s \in \mathbb{C}$) as a way to advance its study.

SIMPLE LEMMA. — *There is a bijection between closed orbits τ for σ^f of least period $\ell(\tau)$ and closed orbits $\{x, \sigma x, \dots, \sigma^{n-1}x\}$, where $f^n(x) := f(x)+f(\sigma x)+\dots+f(\sigma^{n-1}x)$.*



DEFINITION. — We define a zeta function for a (general) flow $\psi_t : \Omega \to \Omega$ by $\zeta(s) = \prod_\tau (1 - e^{-s\ell(\tau)})^{-1}$, ($\tau$ = closed orbit, least period $\ell(\tau)$). *This is a function of* $s \in \mathbb{C}$ (when it converges).

A particular case. — For the suspended flow $\sigma_t^f : X_A^f \to X_A^f$ we have the following expansion :

$$
\begin{aligned}
\zeta_{\sigma^f}(s) &= \prod_\tau (1 - e^{-s\ell(\tau)})^{-1} \\
&= \exp \sum_\tau -\log(1 - e^{-s\ell(\tau)}) && (\exp \circ \log = \text{ identity}) \\
&= \exp \sum_{k=1}^{\infty} \sum_\tau \frac{[e^{-s\ell(\tau)}]^k}{k} && (\log(1-x) = -\sum_{k=1}^{\infty} x^k/k) \\
&= \exp \sum_{k=1}^{\infty} \frac{1}{k} \Big\{ \sum_{n=1}^{\infty} \frac{1}{n} \sum_{\left\{ \substack{\sigma^n x = x \\ n \text{ least}} \right\}} e^{-sf^n(x)\cdot k} \Big\} && (\text{correspondence} \\
& && \gamma \leftrightarrow \{x, \sigma x, \dots, \sigma^{n-1}x\}) \\
&= \exp \sum_{n=1}^{\infty} \frac{1}{n} \sum_{\sigma^n x = x} e^{-sf^m(x)} \qquad (*) && (m = kn,\ k = \text{“iterates”})\,.
\end{aligned}
$$

Another particular case (Geodesic flow). — For the geodesic flow $\phi_t : M \to M$, the closed orbits correspond to (directed) closed geodesics γ of length $\ell(\gamma)$ *i.e.* $\zeta_\phi(s) = \prod_\gamma (1 - e^{-s\ell(\gamma)})^{-1}$. Since our symbolic model is a "good" model for the flow :

PROPOSITION (comparing zeta functions).

$$\zeta_\phi(s) - \zeta_{\sigma^f}(s) = M(s)$$

is analytic and non-zero for $\mathrm{Re}(s) \geqslant h - \varepsilon$ *(for some* $\varepsilon > 0$*).*

Idea of proof. — This is based on the observation that there is "almost" a bijection between closed orbits for ϕ and σ^f. (The difference being of the form $|N(t) - \widetilde{N}(t)| = 0(e^{(h-\varepsilon)t})$).

Note. — A more subtle version of this result (where $M(s)$ is replaced by a product of zeta functions) appears in Bowen, Amer. J. Math., vol. 95 (1973).

K. The domain of the zeta functions

We now want to put together all the preceeding material to construct a meromorphic extension for $\zeta_{\sigma^f}(s)$ (and thus $\zeta_\phi(s)$, using the proposition). The main point is the following :

TECHNICAL LEMMA. — *Assume that we write for* $s = \sigma + it$:

$$\mathcal{L}^n_{-sf} = \sum_{i=1}^{N} \lambda_i^n P_i + \mathcal{U}^{(n)}_{-sr}, \quad \text{with } \varlimsup_n \|\mathcal{U}^{(n)}_{-sr}\|^{1/n} = \rho_s$$

(λ_i *= isolated eigenvalues,* $\mathcal{U}$ *includes the essential spectrum) then we can estimate*

$$\sum_{\sigma^n x = x} e^{-sf^n(x)} = \sum_{i=1}^{N} \lambda_i^n + O(\rho_s)\ . \tag{$*$}$$

Note. — This plays the role of taking a "trace" although $\mathcal{L}_{-sr}$ is *not* trace class. The identity $(*)$ is clearer when we observe that $(\mathcal{L}^n_{-sr} 1)(x) = \sum_{\sigma^n y = x} e^{-sf^n(y)}$.

Recall that ;

$$\begin{aligned}
\zeta_{\sigma^f}(s) &= \exp \sum_{n=1}^{\infty} \frac{1}{n} \sum_{\sigma^n x = x} e^{-sf^n(x)} \\
&= \exp \Big[\sum_{n=1}^{\infty} \frac{1}{n} \big(\sum_{i=1}^{N} \lambda_i^n + O(\rho_s) \big) \Big] \\
&= \exp \Big[\sum_{i=1}^{n} -\log(1 - \lambda_i) + O(\rho_s) \Big]
\end{aligned}$$

$$= \frac{e^{u(s)}}{\prod_{i=1}^{n}(1-\lambda_i(s))}, \tag{**}$$

(where $u(s)$ is analytic, by uniform convergence,, provided $\rho_s < 1$).

This gives us our main result on the domain of zeta functions .

THEOREM (Domain of zeta functions). — *Both* $\zeta_{\sigma f}(s)$ *and* $\zeta_\phi(s)$ *have non-zero meromorphic extensions to half-planes* $\mathrm{Re}(s) \geqslant h - \varepsilon$ *with :*

(i) No poles for $\mathrm{Re}(s) > h$ *;*

(ii) Only a single (simple) pole on $\mathrm{Re}(s) = h$, *at* $s = h$.

The identity (**) tells us that the meromorphic extension is controlled by the spectrum of $\mathcal{L}_{-sf}$ (and, particularly, $\lambda_{-\sigma f} = e^{P(-\sigma f)}$). The following "little lemma" helps us understand this :

LEMMA. — *The map* $\begin{cases} \mathbf{R} \longrightarrow \mathbf{R} \\ t \longmapsto P(-\sigma f) \\ h \longmapsto 0 \end{cases}$ *is surjective (and monotone decreasing). In particular,* $\exists \varepsilon > 0$ *with* $\lambda_{-(h-\varepsilon)f}\theta = 1$.

(The proof is an easy exercise using $f \geqslant 0$ and the variational principle, *cf.* [PP]).

Proof of theorem. — The domain follows from (*), the essential spectral radius of $\mathcal{L}_{-sf}$ being $\lambda_{-\sigma f}\theta \leqslant \rho_s$ (Ruelle operator theorem) and the above lemma.

The pole at $s = h$ (and lack of poles for $\mathrm{Re}(s) > h$) occurs because the eigenvalue identity is satisfied for $\lambda_{-hf} \equiv 1$ (*i.e.* the symbolic characterisation of entropy). The pole is simple because

$$\frac{d}{ds}\lambda_{-sf}\Big|_{s=h} = \frac{d}{d\sigma}e^{P(-\sigma f)}\Big|_{\sigma=h} \neq 0 \quad \text{(By first derivative of pressure)} .$$

Finally, a pole at $s = h + it$ corresponds to $\mathcal{L}_{-sf}k = k$. But compared with $\mathcal{L}_{-hf}1 = 1$ we get $\{f^n(x) \mid \sigma^n x = x\} = a\mathbf{Z}^+$, $a > 0$ (*i.e.* all orbits have lengths which are multiples of a single constant).

Standard application. — Given the above properties of $\zeta_\phi(s)$, there is a "standard" argument (based on the proof of the prime number theorem) which gives $\dfrac{\pi_\phi(t)}{e^{ht}/ht} \longrightarrow 1$ as $t \longrightarrow \infty$.

L. Differentiability of entropy

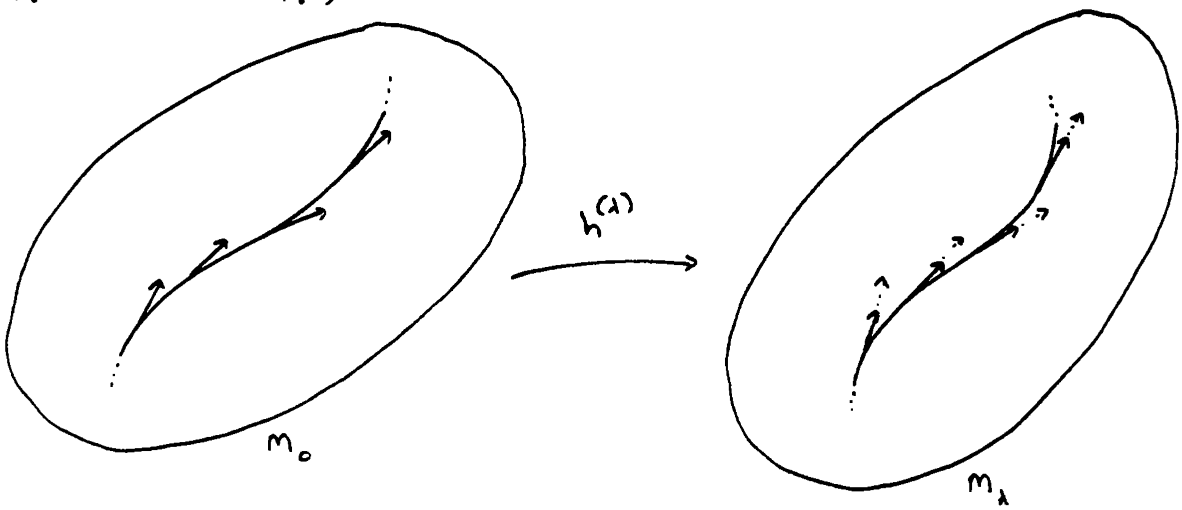
A basic problem is "How does the topological entropy depend on the geometry of V ?".

We need to consider how a small perturbation g_λ in the Riemannian metric g_0 is reflected in the symbolic model.

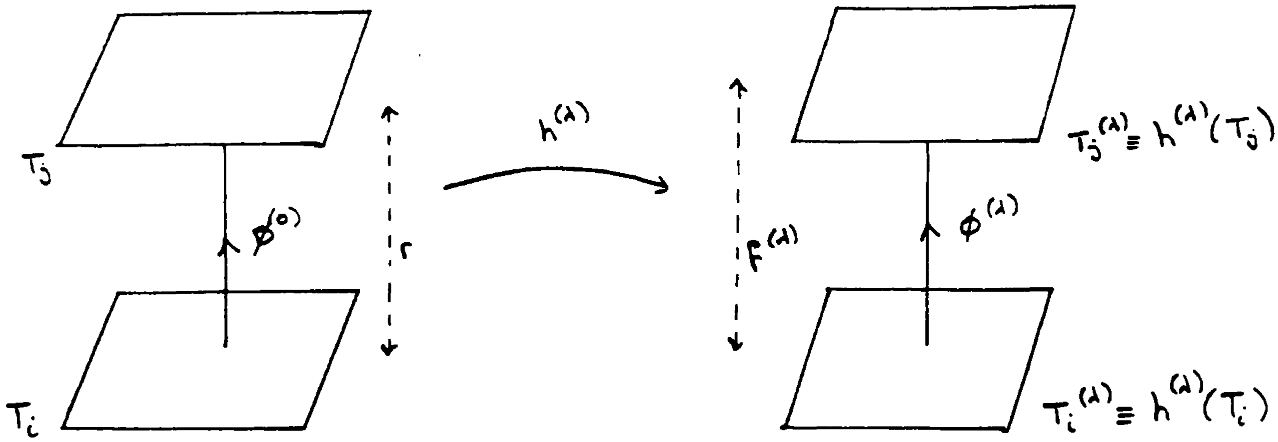
PROPOSITION (structural stability). — *There exists*

(i) homeomorphisms $h^{(\lambda)} : M_0 \to M_\lambda$ $(M_\lambda = T_1(V, g_\lambda))$;

(ii) a "reparameterisation" $\alpha^{(\lambda)} : M_\lambda \to \mathbf{R}$; *such that if we reparameterise* $\phi^{(\lambda)}$ *to* $\psi^{(\lambda)}$ *(i.e.* $X(\psi^{(\lambda)}) = X(\phi^{(\lambda)})\alpha^{(\lambda)} \in \mathcal{X}^0(M)$*) then* $h^{(\lambda)}$ *is a conjugacy (i.e.* $\psi_t^{(\lambda)} \circ h^{(\lambda)} = h^{(\lambda)} \circ \phi_t^{(0)}$*).*



At the symbolic level, we carry over the sections



with differing return times $f^{(\lambda)}(x) = \int_0^{f(x)} \alpha(\phi_t \pi x)\, dt$.

Technical points.

(i) The 2 systems of Markov sections give the same subshift $\sigma : X_A \to X_A$ *but* the different functions $f^{(0)}, f^{(\lambda)} \in C^\alpha(X_A, \mathbf{R})$ (for some $\alpha > 0$).

(ii) The map $\lambda \longrightarrow f^{(\lambda)} \in C^\alpha(X_A, \mathbf{R})$ is C^∞.

THEOREM. — *The topological entropy $\lambda \longmapsto h(\phi^{(\lambda)})$ is C^∞.*

Proof. — By the "symbolic characterisation" we have that $t = h^{(\lambda)}$ is the solution to $P(-tf^{(\lambda)}) = 0$. But we know that $(t, \lambda) \longmapsto P(-tf^{(\lambda)})$ is C^∞ (even C^ω). Thus by the implicit function theorem we get $\lambda \longmapsto h^{(\lambda)}$ is C^∞ (after checking

$$\frac{d}{dt} P(-tf^{(\lambda)})\Big|_{\substack{t=0 \\ \lambda=0}} = -\int f^{(0)}\, d\mu \neq 0\ )$$

(*cf.* Katok, Knieper, Pollicott and Weiss : Invent. Math., 1989).

Mark POLLICOTT
INSTITUT FOURIER
Laboratoire de Mathématiques
URA188 du CNRS
BP 74
38402 St MARTIN D'HÈRES Cedex (France)