Homotopical stability of pseudo-Anosov diffeomorphisms

ALBERT FATHI

The Institute for Advanced Study[†] and CNRS UA 1169, Université Paris-Sud, France

(Received 15 March 1988)

Abstract. We show that the pseudo-Anosov diffeomorphisms have a kind of stability even outside their own homotopy class, this generalizes some results of Lewowicz and Handel. As a corollary, we show that two pseudo-Anosov maps, with the same dilatation coefficient, which are semi-conjugate on the π_1 level are also semiconjugate as dynamical systems by a map which is a ramified cover.

Our main interest in this work is to find when a dynamical system $g: N \rightarrow N$ on a compact connected space can be semi-conjugated onto a pseudo-Anosov map. This is reminiscent of J. Franks work [F], it is also related to work of J. Lewowicz [L] and M. Handel [H1, H2].

A pseudo-Anosov diffeomorphism f of a surface M is a homeomorphism, for which there exists a pair of transverse measured foliations $(\mathcal{F}^s, \mu^s), (\mathcal{F}^u, \mu^u)$ and $\lambda > 1$ such that $f_*(\mathscr{F}^s, \mu^s) = (\mathscr{F}^s, \lambda^{-1}\mu^s)$ and $f_*(\mathscr{F}^u, \mu^u) = (\mathscr{F}^u, \lambda\mu^u)$ - see [FLP] for Thurston's theory of measured foliations and pseudo-Anosov diffeomorphisms. In the beginning, our work was motivated by an attempt to classify pseudo-Anosov maps up to ramified covers, because we wanted to find out the smallest surface on which a given pseudo-Anosov diffeomorphism was 'living' naturally. Our Theorem 2 below shows that this can be reduced to an algebraic semi-conjugacy problem on the fundamental group level. When M. Handel's paper [H1] appeared, it was clear that this problem was related to a form of dynamical 'stability' for pseudo-Anosov diffeomorphisms which is the content of our Theorem 1 below. In fact, the proof of Theorem 1 produces more than its statement, we find a natural hyperbolic extension of the pseudo-Anosov map in which the surface sits naturally as the smallest non-trivial invariant compact connected subset. The space on which the hyperbolic extension exists has a universal cover which is the product of the two trees obtained from the stable and unstable foliations.

The following theorem generalizes some work of J. Lewowicz [L] and M. Handel [H1].

THEOREM 1. Let $f: M \rightarrow M$ be a pseudo-Anosov map of the closed connected surface

[†] Supported by NSF Grant No. DMS-8610730(1).

M. Let $g: N \rightarrow N$ be a homeomorphism of the compact connected space N. Suppose that $\alpha: N \rightarrow M$ is a continuous map such that the diagram:



commutes up to homotopy. If α is not homotopic to a constant, then there exists a closed subset $Y \subset N$, which is g invariant, and a continuous surjective map $\beta : Y \rightarrow M$ such that β is homotopic to $\alpha \mid Y$ and the following diagram commutes:



The second theorem is a generalization of the fact that two homotopic pseudo-Anosov maps are conjugate, it should also be compared with [H2].

THEOREM 2. Let $f: M \to M$ and $g: N \to N$ be pseudo-Anosov maps, with the same dilatation coefficient on the closed connected surfaces M and N. Suppose that $\alpha_*: \pi_1(M) \to \pi_1(N)$ is a non constant algebraic homomorphism such that the following diagram commutes:

$$\begin{array}{ccc} \pi_1(N) & \xrightarrow{g_*} & \pi_1(N) \\ & & & & & \\ & & & & & \\ & & & & \\ & & & & \\ \pi_1(M) & \xrightarrow{f_*} & \pi_1(M), \end{array}$$

where f_* and g_* are respectively representatives of the action of f and g on the fundamental groups. Then there exists a ramified cover $\alpha : N \rightarrow M$ which semi-conjugates g to f, and whose action on fundamental groups can be represented by α_* .

1. Some facts about trees and measured foliations

Let M be a closed connected surface and let \mathscr{F} be a minimal measured foliation on M. Call \tilde{M} the universal cover of M. Denote by $\tilde{\mathscr{F}}$ the pullback of \mathscr{F} to \tilde{M} . We call a leaf of $\tilde{\mathscr{F}}$ either a regular leaf which does not contain a separatrix or the union of a singularity and all separatrices ending in that singularity. Since \mathscr{F} is minimal, any such leaf is a closed subset of \tilde{M} . We call \mathscr{T} the set of leaves $\tilde{M}/\tilde{\mathscr{F}}$. If we define the distance between two leaves of $\tilde{\mathscr{F}}$ as the minimum of the transverse measures of arcs joining the two leaves, we obtain a distance on \mathscr{T} which turns \mathscr{T} into a tree – see Morgan and Shalen [MS, § 2]. If the genus of M is ≥ 2 , then \mathscr{T} is not complete for this distance. We will denote by $\hat{\mathscr{T}}$ its completion, it is also a tree, see [MS, proof of Theorem 2.1.9]. To simplify notations, we will denote by Γ the fundamental group of M the group Γ acts in a natural way as a group of isometries of \mathcal{T} and hence it also acts as a group of isometries of $\hat{\mathcal{T}}$.

LEMMA 1.1. If a is a closed arc contained in $\hat{\mathcal{T}}$ then a minus its endpoints is contained in \mathcal{T} . If $f \in \mathcal{T}$, the set $\hat{\mathcal{T}} \setminus \Gamma f$ is totally disconnected.

In order to prove 1.1, we need a couple of sublemmas.

SUBLEMMA 1.2. Let α (resp. α') be an arc between x and y (resp. x' and y') in a tree. The intersection $\alpha \cap \alpha'$ is an arc of length $l(\alpha \cap \alpha')$ satisfying:

$$l(\alpha \cap \alpha') \geq l(\alpha) - d(x, x') - d(y, y').$$

The proof of this sublemma can be found in [MS, proof of Corollary 2.1.7].

SUBLEMMA 1.3. Let $\hat{\mathcal{T}}$ be the completion of the tree \mathcal{T} . If α is an arc in $\hat{\mathcal{T}}$, then α minus its extremities is contained in \mathcal{T} .

Proof. Call x and y the extremities of α . Choose $x_n \to x$ and $y_n \to y$ with $x_n, y_n \in \mathcal{T}$. Call $\alpha_n \subset \mathcal{T}$ the arc between x_n and y_n . We have $\alpha \cap \alpha_n \subset \mathcal{T}$ and $l(\alpha \cap \alpha_n) \to l(\alpha)$ by 1.2. The sublemma follows easily.

Proof of 1.1. Since the foliation \mathscr{F} is minimal and any non trivial arc in \mathscr{T} contains the image under the map $\tilde{M} \to \mathscr{T}$ of a non trivial arc transverse to $\tilde{\mathscr{F}}$, we obtain that any non trivial arc of \mathscr{T} contains a point of Γf . By 1.2, the same is true for any non trivial arc in $\hat{\mathscr{T}}$. Lemma 1.1 follows from this last fact.

2. Embedding a pseudo-Anosov map in a hyperbolic dynamical system

Let $f: M \to M$ be a pseudo-Anosov map. Let $\lambda > 1$ be its dilatation coefficient and let \mathscr{F}^s and \mathscr{F}^u be its stable and unstable foliations. Call (\mathscr{T}^u, d^u) and (\mathscr{T}^s, d^s) the trees $\tilde{M}/\tilde{\mathscr{F}}^s$ and $\tilde{M}/\tilde{\mathscr{F}}^u$ with their respective metrics. Fix a lift \tilde{f} of f to \tilde{M} . This \tilde{f} induces actions \tilde{F}^u and \tilde{F}^s on the trees \mathscr{T}^u and \mathscr{T}^s which satisfy:

$$\forall a, a' \in \mathcal{T}^u, \qquad d^u(\tilde{F}^u(a), \tilde{F}^u(a')) = \lambda d^u(a, a')$$

$$\forall b, b' \in \mathcal{T}^s, \qquad d^s(\tilde{F}^s(b), \tilde{F}^s(b')) = \lambda^{-1} d^s(b, b').$$

This implies that these actions extend to the completions $\hat{\mathcal{T}}^{\mu}$ and $\hat{\mathcal{T}}^{s}$ and that these extensions verify the same equalities.

If we look at the product action $\tilde{F} = \tilde{F}^s \times \tilde{F}^u$ on $\tilde{Z} = \hat{\mathcal{T}}^s \times \hat{\mathcal{T}}^u$ with the product metric $d = d^s + d^u$ we obtain what we can call a metrically split hyperbolic homeomorphism – generalizing to arbitrary metric spaces [F, Definition on p. 67].

There is an inclusion $\tilde{M} \to \tilde{Z}$, which is obtained by sending a point in M to the pair of leaves - one from $\tilde{\mathscr{F}}^s$ and the other from $\tilde{\mathscr{F}}^u$ - that contain it. It is well known that the metric d induces on \tilde{M} the usual topology.

The action of $\Gamma = \pi_1(M)$ on \tilde{M} induces isometries on \tilde{Z} in a natural way. The inclusion $\tilde{M} \hookrightarrow \tilde{Z}$ is equivariant with respect to these actions of Γ .

If we write $\tilde{f}\gamma = f_{\#}(\gamma)\tilde{f}$ for $\gamma \in \Gamma$, we also have $\tilde{F}\gamma = f_{\#}(\gamma)\tilde{F}$.

LEMMA 2.1. There exists $\varepsilon > 0$ such that:

 $\forall \gamma \in \Gamma \setminus \{ \mathrm{Id} \}, \quad \forall z \in \tilde{Z}, \, d(\gamma z, z) \geq \varepsilon.$

Proof. If $z = (a^u, a^s) \in \tilde{\mathcal{F}}^u \times \tilde{\mathcal{F}}^s$ then $\gamma z = (\gamma a^u, \gamma a^s)$. This implies:

$$(\gamma z, z) = d^{u}(\gamma a^{u}, a^{u}) + d^{s}(\gamma a^{s}, a^{s})$$
$$\geq i(\gamma, \mathcal{F}^{u}) + i(\gamma, \mathcal{F}^{s}).$$

It is well known that there exists $\varepsilon > 0$ such that:

d

$$\forall \gamma \in \Gamma \setminus \{ \mathrm{Id} \}, \, i(\gamma, \mathcal{F}^u) + i(\gamma, \mathcal{F}^s) \geq \varepsilon. \qquad \Box$$

COROLLARY 2.2. The action of Γ on \tilde{Z} is properly discontinuous.

If we call $Z = \tilde{Z}/\Gamma$, we obtain a metric space since d is Γ equivariant. The map $\tilde{Z} \to Z$ appears as the universal cover of Z since \tilde{Z} is contractible and locally contractible. The inclusion $\tilde{M} \to \tilde{Z}$ gives an inclusion $M \to Z$ which is an isomorphism on the fundamental group. The map \tilde{F} gives a map $F: Z \to Z$ which is a hyperbolic homeomorphism whose universal cover is metrically split. The images of the foliations $\{\hat{\mathcal{T}}^s \times b | b \in \hat{\mathcal{T}}^u\}$ and $\{a \times \hat{\mathcal{T}}^u | a \in \hat{\mathcal{T}}^s\}$ under the map $\tilde{Z} \to Z$ give the stable and unstable foliations of F. The restriction of F to M is f.

LEMMA 2.3. In the situation described above, two continuous maps $g_1, g_2: X \rightarrow M$ are homotopic as maps with values in M if and only if they are homotopic as maps with values in Z.

Proof. Since the genus of the surface M is ≥ 1 , we can endow M with a Riemannian metric such that any pair of points in the universal cover \tilde{M} can be connected by a unique geodesic. Given any path $\alpha \subset Z$ connecting the two points $x, y \in M$, a lift $\tilde{\alpha}$ to \tilde{Z} connects two points $\tilde{x}, \tilde{y} \in \tilde{M}$, call α' the image in M of the unique geodesic connecting \tilde{x}, \tilde{y} in \tilde{M} . The map $\alpha \mapsto \alpha'$ is well defined and continuous in the compact open topology. The lemma follows routinely from this fact.

LEMMA 2.4. If $x \in M$, then $Z \setminus (W^s(x, F) \cup W^u(x, F))$ is totally disconnected.

Proof. Choose a point $(f, f') \in \tilde{Z}$ above x. We have $f \in \mathcal{T}^s$, $f' \in \mathcal{T}^u$. It follows from 1.1 that $(\hat{\mathcal{T}}^s \setminus \Gamma f) \times (\hat{\mathcal{T}}^u \setminus \Gamma f')$ is totally disconnected. But this product is precisely the inverse image of $Z \setminus (W^s(x, F) \cup W^u(x, F))$ under the covering map $\tilde{Z} \to Z$.

PROPOSITION 2.5. If X is a closed connected non empty subset of Z which is F invariant then either it is reduced to a point or it contains M.

Proof. Suppose that X is not reduced to a point. Since the periodic points of f = F | M are dense in M and X is closed, it suffices to show that X contains these periodic points. Fix such a periodic point p. Since X is closed and invariant under F, it suffices to show that either $W^s(p, F)$ or $W^u(p, F)$ intersects X. But this follows clearly from 2.4, since X is connected and not reduced to a point.

The next theorem is a generalization of Theorem 1.

290

THEOREM 2.6. Suppose that we have a diagram:

$$\begin{array}{ccc} N & \xrightarrow{g} & N \\ & \alpha & & & \downarrow^{\alpha} \\ z & \xrightarrow{F} & z \end{array}$$

commutative up to homotopy, where g is a homeomorphism of the compact connected space N. Then there exists a continuous map $\beta : N \rightarrow Z$, homotopic to α , and such that the diagram:



commutes. Moreover, if α is not homotopic to a constant, then the image $\beta(N)$ contains M. It follows that $Y = \beta^{-1}(M)$ is invariant under g and that $\beta \mid Y$ is a continuous surjection onto M, which is a semi-conjugacy between $g \mid Y$ and f. Moreover, if $\alpha(Y) \subset M$ then $\alpha \mid Y$ and $\beta \mid Y$ are homotopic as maps with values in M.

Proof. Since Z has a universal cover on which F has a hyperbolic metrically split lift with a complete metric. It is easy to see that the machinery developed by J. Franks in [F, § 4] can be applied to give $\beta : N \rightarrow Z$, homotopic to α which gives a semi-conjugacy:

If α is not homotopic to a constant, by 2.3, the image $\beta(N)$ is a compact connected F invariant subset of Z which is not reduced to a point. By 2.5, we have $M \subset \beta(N)$. The last assertion follows from 2.3.

3. Pseudo-Anosov diffeomorphisms with the same dilatation coefficient The following lemma is certainly well known.

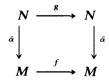
LEMMA 3.1. Let $g: N \rightarrow N$ be a pseudo-Anosov diffeomorphism on a closed connected surface, with dilatation coefficient $\lambda > 1$. Suppose that $Y \subset N$ is a compact g-invariant subset, if the topological entropy of $g \mid Y$ is $\log \lambda$, then Y = N.

Proof. From [FLP, Exposé 10 § IV and § VI], there exists an irreducible subshift of finite type (Σ_A, σ_A) , whose topological entropy is $\log \lambda$, and a surjective semiconjugacy $\theta: \Sigma_A \to N$ between σ_A and g. The closed subset $\theta^{-1}(Y)$ is σ_A -invariant and the topological entropy of σ_A restricted to that subset is $\log \lambda$, it follows from [CP, Theorem 3.3] that $\theta^{-1}(Y) = \Sigma_A$. The surjectivity of θ implies that Y = N.

Let $f: M \to M$ and $g: N \to N$ be pseudo-Anosov maps, with the same dilatation coefficient λ on the closed connected surfaces M and N. Suppose that $\alpha_*: \pi_1(N) \to \pi_1(M)$ is a non constant algebraic homomorphism such that the following diagram commutes:

where f_* and g_* respectively represent f and g on the fundamental group.

The first remark is that there exists a map $\bar{\alpha}: N \to M$, which can be represented by α_* on π_1 and such that the diagram:



commutes up to homotopy. These are standard facts from algebraic topology. By Theorem 2.6, there exists $Y \subset N$ a g-invariant subset and a surjective semi-conjugacy $\alpha: Y \rightarrow M$, between $g \mid Y$ and f, with α homotopic to $\overline{\alpha} \mid Y$. Since the topological entropy of both f and g is $\log \lambda$, the topological entropy of $g \mid Y$ is also $\log \lambda$. By 3.1, we have Y = N and, in fact, the map α is a surjective semi-conjugacy between g and f, which is homotopic to $\overline{\alpha}$. The following lemma finishes the proof of Theorem 2.

LEMMA 3.2. Any non constant semi-conjugacy between two pseudo-Anosov maps, with the same dilatation coefficient on closed connected surfaces is a ramified cover.

Proof. Let us denote by $\varphi: N \to M$ a non constant semi-conjugacy between the two pseudo-Anosov diffeomorphisms $g: N \to N$ and $f: M \to M$ with the same dilatation coefficient λ . We denote by (\mathcal{F}^s, μ^s) , (\mathcal{F}^u, μ^u) the invariant stable and unstable measured foliations of f and by (\mathcal{G}^s, ν^s) , (\mathcal{G}^u, ν^u) those of g. Since φ is a semi-conjugacy it takes a leaf of \mathcal{G}^s (resp. \mathcal{G}^u) to a leaf of \mathcal{F}^s (resp. \mathcal{F}^u).

What we mean by a segment contained in a \mathscr{G}^{u} leaf is a subset of a \mathscr{G}^{u} leaf which is homeomorphic to [0, 1] and if it contains a singularity then it must be entirely contained in the union of the singularity and of two separatrices adjacent to the same sector. Remark that the holonomy along \mathscr{G}^{s} is always defined on one side of such a segment. If α is a segment in a \mathscr{G}^{u} leaf, we denote by $v^{s}(\alpha)$ its v^{s} measure and by $\overline{\mu}^{s}(\alpha)$ the measure of the \mathscr{F}^{u} segment which has the same extremities as $\varphi(\alpha)$. It is clear that $\overline{\mu}^{s}$ is continuous on its domain of definition endowed with the compact open topology. Since g and f have the same dilatation coefficient λ , we have:

$$\forall \alpha \mathcal{G}^{u} \text{-segment } \bar{\mu}^{s}(f(\alpha)) = \lambda \bar{\mu}^{s}(\alpha).$$
(1)

If γ is a segment of a \mathscr{G}^{μ} leaf which is the union of the two subsegments α and β

then we have:

$$\bar{\mu}^{s}(\gamma) \leq \bar{\mu}^{s}(\alpha) + \bar{\mu}^{s}(\beta).$$
⁽²⁾

If we move a segment in a \mathscr{G}^{u} leaf by holonomy along \mathscr{G}^{s} leaves to another segment in a \mathscr{G}^{u} leaf, the two segments will have the same value under $\bar{\mu}^{s}$.

Our goal is to show that there exists $\rho > 0$ such that for any segment α in a \mathscr{G}^{u} leaf $\bar{\mu}^{s}(\alpha) = \rho \nu^{s}(\alpha)$. We will use the fact that g has a Markov partition $\mathscr{R} = \{R_{1}, \ldots, R_{n}\}$ - see [FLP, Exposé 10 § IV, § V and § VI] for the definition and properties of Markov partitions.

Denote by $S_i^u \subset R_i$ any segment in a \mathscr{G}^u leaf which goes across R_i . The values $a_i = \nu^s(S_i^u)$ and $b_i = \overline{\mu}^s(S_i^u)$ do not depend on the choice of S_i^u because any two such choices differ by holonomy along \mathscr{G}^s leaves. Call x_{ij} the number of times $g^{-1}(R_j)$ crosses the interior of R_i . This gives a positive matrix $X = (x_{ij})_{1 \le i, j \le n}$. From [FLP, Exposé 10 Lemma 1 p. 205], the matrix X has power which is strictly positive. Moreover, if $X^l = (x_{ij}^{(l)})_{1 \le i, j \le n}$, then $(x_{ij}^{(l)})_{1 \le i, j \le n}$ is the number of times that $f^{-l}(R_j)$ crosses R_i . It follows from this remark and (2) above that, if we define A (resp. B) as the vector with components a_1, \ldots, a_n (resp. b_1, \ldots, b_n), then we have:

$$\forall l \ge 1, \, \lambda^{l} A = X^{l} A \quad \text{and} \quad \lambda^{l} B \le X^{l} B. \tag{3}$$

Call $\tilde{a}_1, \ldots, \tilde{a}_n$ the components of the eigenvector with eigenvalue λ of the transpose of X normalized by the relation $\sum_{i=1}^{n} \tilde{a}_i a_i = 1$. From Perron-Frobenius theory, it is well known that $\tilde{a}_i > 0$ and $\lim_{i \to \infty} \lambda^{-i} X^i B = \rho A$, with $\rho = \sum_{i=1}^{n} \tilde{a}_i b_i$. It follows from (3) that $B \le \rho A$. If one of the components of the last inequality was strict multiplying the *i*-component by the strictly positive \tilde{a}_i and summing would give $\rho = \sum_{i=1}^{n} \tilde{a}_i b_i < \rho \sum_{i=1}^{n} \tilde{a}_i a_i = \rho$ which is impossible. So we have:

$$\forall i=1,\ldots,n, \qquad b_i=\rho a_i. \tag{4}$$

Suppose now that α is a segment in a \mathscr{G}^{u} leaf. For each j = 1, ..., n and each $l \ge 1$, call α_{j}^{l} the number of \mathscr{G}^{u} segments intersecting α which are contained in $f^{-l}(R_{i})$ and go across it. Using (1) through (4), we obtain:

$$\bar{\mu}^{s}(\alpha) \leq \sum_{i=1}^{n} \alpha_{j}^{l} \lambda^{-l} b_{j} = \sum_{i=1}^{n} \alpha_{j}^{l} \rho \lambda^{-l} a_{j} \rightarrow \rho \nu^{s}(\alpha).$$

This shows the inequality $\bar{\mu}^{s}(\alpha) \leq \rho \nu^{s}(\alpha)$. Note now that ρ cannot be zero, because that would imply that the image of $\varphi(N)$ would be contained in one leaf of \mathscr{F}^{s} and hence, by the g invariance, it would be reduced to one point. Suppose that some segment β in a \mathscr{G}^{u} leaf verifies $\bar{\mu}^{s}(\beta) < \rho \nu^{s}(\beta)$. Choose a point x in the interior of β and call x_{∞} an accumulation point of the sequence $g^{-l}(x)$, $l \geq 0$. It is clear that the unstable leaf through x_{∞} contains arbitrarily small segments α , containing x_{∞} , and such that $\bar{\mu}^{s}(\alpha) < \rho \nu^{s}(\alpha)$. The density of the stable leaf through x_{∞} and the invariance by holonomy show that the same strict inequality is true for any segment in a \mathscr{G}^{u} leaf. This is absurd, since we have equality for any segment in a \mathscr{G}^{u} leaf contained in a R_{i} and going across it.

By rescaling ν^s , we can assume that $\rho = 1$. This allows us to interpret what we obtained in the following way: the map φ is an isometry of any non singular leaf

of \mathscr{G}^{u} endowed with the metric defined by ν^{s} onto the corresponding leaf of \mathscr{F}^{u} endowed with the metric defined by μ^{s} . Of course, the same result can be obtained with stable foliations. It is now easy to finish the proof of the lemma.

Acknowledgements. I am most grateful to Mike Handel not only for his active help in writing this paper, but also for the constant stimulation provided by his work. My debt to his papers [H1, H2] is enormous. The proof I am giving for Theorem 1 was found in Spring 1986, while I was enjoying the hospitality of the Mathematics Institute at the University of Warwick for the 'Special year in smooth ergodic theory'.

REFERENCES

- [CP] E. M. Coven & M. E. Paul. Endomorphisms of irreducible subshifts of finite type. Math. Sys. Th. 8 (1974-75), 167-175.
- [FLP] A. Fathi, F. Laudenbach & V. Poenaru. Travaux de Thurston sur les surfaces. Astérisque 66–67 (1979).
- [F] J. Franks. Anosov diffeomorphisms. Proc. Symp. in Pure Mathematics 14 (1970), 61-94.
- [H1] M. Handel. Global shadowing of pseudo-Anosov homeomorphisms. Ergod. Th. & Dynam. Sys. 5 (1985), 373-377.
- [H2] M. Handel. Entropy and semi-conjugacy in dimension two. Ergod. Th. & Dynam. Sys. 8 (1988), 585-596.
- [L] J. Lewowicz. Persistence in expansive systems. Ergod. Th. & Dynam. Sys. 3 (1983), 567-578.
- [MS] J. Morgan & P. B. Shalen. Valuations, trees, and degenerations of hyperbolic structures: I. Ann. Math. 120 (1984), 401-476.