Nova Série
BOLETIM
DA SOCIEDADE BRASILEIRA DE MATEMÁTICA

Bol. Soc. Bras. Mat., Vol. 20, No. 1, 113-133

Expansive homeomorphisms of surfaces

Jorge Lewowicz

Abstract. Let f be an expansive homeomorphism of a compact oriented surface M. We show that $\mathbf{S}^2$ does not support such an f, and that f is conjugate to an Anosov diffeomorphism if $M = \mathbf{T}^2$, and to a pseudo-Anosov map if M has genus ≥ 2. These results are consequences of our description of local stable (unstable) sets: every $x \in M$ has a local stable (unstable) set that consists of the union of r arcs that meet only at x. For each $x \in M$ $r = 2$, except for a finite number of points, where $r \geq 3$.

Introduction

In this paper we study the topological classification of expansive homeomorphisms of compact orientable surfaces without boundary.

Let M be a compact metric space; a homeomorphism $f: M \to M$ is expansive if there exists $\alpha > 0$ such that if $x, y \in M$, $x \neq y$, then $\text{dist}(f^n(x), f^n(y)) > \alpha$ for some $n \in \mathbf{Z}$. The homeomorphisms f, g are topologically equivalent or conjugate if there exists another homeomorphism $h: M \to M$ such that $h \circ f = g \circ h$; obviously, expansivity is a conjugacy invariant. Anosov diffeomorphisms, pseudo-Anosov maps and their smooth models [2], [11], [4], [9], are examples of expansive homeomorphisms of compact manifolds; the subshifts in $K^{\mathbf{Z}}$ (K a finite set) are also expansive.

As it is well known, there are no expansive homeomorphisms of $\mathbf{S}^1$; on the other hand every compact orientable surface of genus ≥ 1 supports such a homeomorphism [10]. We prove (Section 4) that there are no expansive homeomorphisms of $\mathbf{S}^2$.

With respect to topological equivalence of expansive diffeomorphisms, in [3] it is shown that an Anosov diffeomorphism on the torus $\mathbf{T}^2$ is conjugate to a linear hyperbolic isomorphism. We show (section 5) that every expansive homeomorphism of $\mathbf{T}^2$ is also conjugate to such a linear map, and that on compact orientable surfaces of genus ≥ 2, an expansive homeomorphism is topologically equivalent to a pseudo-Anosov map (section 6) (I learned that K. Hiraide has

Received 22 April 1989.

obtained the same results). This implies, in particular that, roughly, the dynamics of an expansive homeomorphism of a surface is preserved under C^0-perturbations: expansive homeomorphisms are topologically stable on $\mathbf{T}^2$ and persistent [8] on other surfaces.

In the proof of the above mentioned results of [3] concerning Anosov diffeomorphisms f, the existence of continuous foliations of stable and unstable manifolds plays a fundamental role. The stable (unstable) set of x consists of those y such that $\mathrm{dist}(f^n(x), f^n(y)) \to 0$ for $n \to \infty$ (resp. $n \to -\infty$). The fact that these sets are differentiable manifolds at any x is a consequence of the rich uniform behaviour of the tangent map of Anosov diffeomorphisms, which does not hold, in general, even for smooth expansive diffeomorphisms. On the other hand, expansivity means, from the topological point of view, that any point of the space M has a distinctive dynamical behaviour. Therefore a stronger interaction between the topology on M and the dynamics could be expected. For instance, while for subshifts on compact perfect spaces, the stable set of some point may reduce to the point, for expansive homeomorphisms of compact manifolds, this is never the case. Moreover, any sufficiently small closed neighbourhood of each point $x \in M$ contains a connected subset of the stable (unstable) set of x, that joins x to the boundary of the neighbourhood (Lemma 2.1). When $\dim M = 2$, since such a connected stable piece meets an ustable one at one point, we prove, using the nice topological properties of the plane, that they are locally connected (see Lemma 2.3). Thus, such a piece contains arcs that join any pair of its points; from this fact we obtain that, locally, the stable (unstable) set of x is the union of p arcs, $p \geq 2$, which meet only at x and separate the sectors limited by the unstable (stable) arcs (Proposition 3.6). Furthermore, every x has a neighbourhood that consists of points at which, except perhaps x itself, $p = 2$; therefore $p > 2$ only on a finite set. On the complement of this set we get continuous transversal foliations of stable and unstable manifolds.

With arguments analogous to those used in the classical Poincaré- Bendixson's theory, we show, essentially, that the lifting of a stable (unstable) manifold to the universal covering of M is a closed submanifold (see Lemma 4.1); this implies readily that $\mathbf{S}^2$ does not support expansive homeomorphisms.

The classification of expansive homeomorphisms of $\mathbf{T}^2$ follows from the previous conclusions and from the results of [3] on semiconjugacy: if g is a homeomorphism of $\mathbf{T}^2$ homotopic to a linear hyperbolic isomorphism f, there exists a continuous surjective $h: \mathbf{T}^2 \to \mathbf{T}^2$ such that $h \circ g = f \circ h$ (g is *semiconjugate* to f). The fact that the lifting of h to $\mathbf{R}^2$ is a proper map, together with the mentioned properties of liftings of stable (unstable) manifolds allow to prove that, for expansive g, h is also injective. Since some diameters grow exponentially

under the iterates of a lifting of g, we get that g is homotopic to f (Lemma 5.4).

For surfaces M with higher genus, the results on semiconjugacy of [5] state that, if g is homotopic to a pseudo-Anosov f, there exists a g-invariant compact set $C \subset M$ such that $g|C$ is semiconjugate to f. We show that when g is expansive, $C = M$ (see Lemma 6.2). The classification of expansive homeomorphisms of M follows now from arguments similar to those used in the case of the torus and from Thurston's isotopy classification theorem [2], [11].

1. Lyapunov functions

Consider an expansive homeomorphism f of a compact connected manifold M, endowed with some riemannian metric. Let $\alpha > 0$ be an expansivity constant for f, i.e., $\mathrm{dist}\,(f^n(x), f^n(y)) \leq \alpha$ for every $n \in \mathbf{Z}$ implies $x = y$. Call $(\hat{M}, \phi)$ the suspension of f, $p: M \times \mathbf{R} \to M$ the canonical projection, and M_t the manifold $p(M \times \{t\})$, $t \in \mathbf{R}$; we shall identify M to M_0.

Let V be a real continuous function defined on a neighbourhood of the diagonal of $N = \cup_{t \in \mathbf{R}} M_t \times M_t$ and such that $V(\xi, \xi) = 0$ for $\xi \in \hat{M}$. Define $\dot{V}$, the derivative of V along the flow, as

$$\dot{V}(\xi, \eta) = \lim_{t \to 0} \frac{1}{t} \left(V(\phi(\xi, t), \phi(\eta, t) - V(\xi, \eta)\right).$$

Analogously, define $\ddot{V}$ by replacing V by $\dot{V}$ in the above limit. When $V(\xi,\eta) > 0$ and $\ddot{V}(\xi, \eta) > 0$ for $\xi \neq \eta$, we call V a Lyapunov function for the suspension flow of f (see [7], p. 197). An idea about the geometric meaning of the existence of a Lyapunov function for such a flow may be obtained as follows. Let k be a small positive number, $\xi \in \hat{M}$ and $K_t = \{\eta \in M_t : V(\phi(\xi, t), \eta) \leq k\}$. We may think of $\cup_{t \in \mathbf{R}} K_t$ as a tube with axis $\phi(\xi, t)$, $t \in \mathbf{R}$. The fact that $\ddot{V} > 0$ implies that if a trajectory of the suspension flow is in the border of the tube for $t = t_0$, then, it stays in the tube in no neighbourhood of t_0.

In [7], section 4, the arguments previous to proposition 4.3 show that if f is an expansive diffeomorphism, there exists a Lyapunov function for its suspension flow. The fact that f is a diffeomorphism is only used to obtain the flow ψ defined on N by the vector field aX, where X is the vector field on N defined by the restriction to N of $\phi \times \phi$ (that we shall call also ϕ), and a is a smooth real function that vanishes on a compact subset A of N and is positive on $N - A$. Since we will use Lyapunov functions for the suspension flow of an expansive homeomorphism, we include here some simple remarks that allow the use of the arguments of [7], section 4, to the case of homeomorphisms.

Let N be a compact manifold and A a compact subset of N. Let $\phi: N \times \mathbf{R} \to N$ be a dynamical system, and for $x \in A$, let $\sigma(t) = \sup_{x \in A} a(\phi(x, t))$, where

$a: N \to \mathbf{R}$ is a continuous function that vanishes on A and is positive elsewhere. Obviously, σ is continuous and $\lim_{t\to 0}\sigma(t) = 0$.

Lemma 1.1. *There exists a continuous increasing $h: \mathbf{R} \to \mathbf{R}$, $h(0) = 0$, such that*

$$\int_0 (h(\sigma(t))^{-1} dt = \infty.$$

Proof. Let a_n be a decreasing sequence $a_n \to 0$, such that for $|t| \le a_n$, $\sigma(t) \le 1/n$, and $b_n \to 0$ be another decreasing sequence such that $\sum_{n=1}^{\infty}(a_{n+1}-a_n)/b_n = \infty$. Pick a continuous increasing $h, h(0) = 0$, so that $|h(t)| \le b_n$ if $1/(n+1) \le t \le 1/n$; then $\int_0 (h(\sigma(t))^{-1}dt \ge \sum_{n=1}^{\infty}(a_{n+1} - a_n)/b_n$.

Clearly, such an h may be taken to be smooth, $h'(t) > 0$ if $t \ne 0$, and $h(1) = 1$. Put $g = h \circ a$; then g vanishes on A, is positive on $N - A$, $g = 1$ if $a = 1$, and $\int_0 (g(\phi(x,t))^{-1}dt = \infty$ for $x \in A$. Now, let us return to our compact manifold $N = \cup_{t\in\mathbf{R}} M_t \times M_t$, endowed with some metric, and to its flow ϕ defined, for $x = (\xi,\eta) \in N$, by $\phi(x,t) = (\phi(\xi,t), \phi(\eta,t))$, as mentioned above. Let σ, δ and a be defined as in [7], p. 201; this time $a: N \to \mathbf{R}$ is only continuous. For $x \in N, x \notin A = \{(\xi,\eta) \in N : \mathrm{dist}(\xi,\eta) \ge \delta\}$ let $\psi(x,t) = \phi(x,\tau(t))$, where $\tau(t) = \tau_x(t)$ is the inverse function of $t = \int_0^{\tau} g(\phi(x,s))^{-1}ds$, and for $x \in A$, let $\psi(x,t) = x$.

Lemma 1.2. *ψ is a flow on N.*

Proof. We show first that for $x \notin A$, ψ is defined for every $t \in \mathbf{R}$. For such an x, either $\phi(x,\tau^*) \in A$ for some $\tau^* > 0$, or, because of expansivity, $\mathrm{dist}(\phi(x,t), A) > \lambda$ for some $\lambda > 0$ and each $t > 0$. Thus, in the second case, $\psi(x,t)$ is defined for all $t > 0$; with respect to the first one, let τ^* be the first positive number for which $\phi(x,t) \in A$. Then, $\int_0^{\tau}(g(\phi(x,s))^{-1} \to \infty$ when $\tau \to \tau^*$ as may easily be shown through the change of variables $s = u + \tau^*$ on account of the previous lemma and the fact that $\phi(x,\tau^*) \in A$. Therefore $\psi(x,t)$ is defined for all $t \ge 0$; the proof for $t \le 0$ is similar. We finish the proof of the lemma by showing that ψ is continuous. If $x \notin A$, the continuity follows from that of ϕ and g since $g(\psi(x,t)) > 0$ for $t \in \mathbf{R}$; for $x \in A$ let x_n, t_n be such that $x_n \to x$, $t_n \to t$, $x_n \notin A$, and, say, $t_n > 0$, $n = 1, 2, \ldots$.

Since $\int_0^{\tau(t_n)}(g(\phi(x_n,s))^{-1}ds) = t_n > C\tau(t_n)$ for some positive constant C, we have, by taking convergent subsequences, that $\tau(t_n) \to \tau^*$, $0 \le \tau^* < \infty$. As $g(\phi(x_n,s))$ converges to $g(\phi(x,s))$ uniformly on $[0,\tau^*]$, we must have $\tau^* = 0$. Thus, $\psi(x_n,t_n)$ converges to $\phi(x,0)$.

Let f be an homeomorphism of the compact manifold M, and $\mathbf{V}: M \times M \to \mathbf{R}$ a continuous function that vanishes on the diagonal; for $\xi, \eta \in M$, call $\Delta\mathbf{V}$ the difference, $\Delta\mathbf{V}(\xi,\eta) = \mathbf{V}(f(\xi), f(\eta)) - \mathbf{V}(\xi,\eta)$. We say that $\mathbf{V}$ is a Lyapunov function for f if $\mathbf{V}$ and $\Delta(\Delta\mathbf{V})$ are positive for (ξ,η) in a neighbourhood of the diagonal, $\xi \ne \eta$.

Theorem 1.3. *Let f be a homeomorphism of the compact manifold M. The following assertions are equivalent*

i) *f is expansive,*

ii) *There exists a Lyapunov function for the suspension flow of f.*

iii) *There exists a Lyapunov function for f.*

Proof. The equivalence of i) and ii) follows, on account of the previous lemmas, from the already mentioned arguments in [7], section 4, and from Lemma 3.3 in the same paper. To construct a Lyapunov function for an expansive f, let V be a Lyapunov function for its suspension flow; identify M_0 with M and define $\mathbf{V}(\xi,\eta) = V(\xi,\eta)$.

Since we shall use, for homeomorphisms, Lemma 2.7 of [8], we recall that $x \in M$ is a stable (unstable) point of the homeomorphism f of the metric space M if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $y \in M$ and $\mathrm{dist}(x,y) < \delta$ imply $\mathrm{dist}(f^n(x), f^n(y)) < \varepsilon$ for $n \ge 0$ (resp. $n \le 0$).

Lemma 1.4. *An expansive homeomorphism of a compact manifold has no stable (unstable) points.*

Proof. See [8], p. 573.

2. Stable and unstable sets

Let M be a riemannian manifold and f an expansive homeomorphism of M, with expansivity constant $\alpha > 0$. For $\delta > 0$, and $x \in M$, let $B_\delta(x) = \{y \in M : \mathrm{dist}(x,y) < \delta\}$; choose $\delta_1, \delta_2, k, 0 < \delta_1 < \delta_2 < \alpha$, $k > 0$ in such a way that for every $x \in M$,

$$B_{\delta_1}(x) \subset \{y \in M : V(x,y) \le k\} \subset B_{\delta_2}(x);$$

here, as in the previous section, V denotes a Lyapunov function for the suspension flow of f. Also, we keep with the identification of M to M_0.

Lemma 2.1. *Let $A \subset M$ be an open set, $x \in A \subset B_{\delta_1}(x)$. Then there exists a compact connected set C, $x \in C \subset \mathrm{clos}(A)$, $C \cap \partial A \ne \varnothing$, such that for every $y \in C$, $\mathrm{dist}(f^n(x), f^n(y)) \le \delta_2$, if $n \ge 0$.*

Proof. Assume this is not the case. Then there exists $N > 0$ so that for every compact connected $D \subset \mathrm{clos}(A)$ and joining x to ∂A, there exists $z \in D$ and

$n, 0 \le n \le N$, such that $\operatorname{dist}(f^n(x), f^n(y)) > \delta_2$. Otherwise, for all $n \ge 0$, we could find a compact connected set $D_n \subset \operatorname{clos}(A)$, joining x to ∂A and such that for every $y \in D_n$, $\operatorname{dist}(f^m(x), f^m(y)) \le \delta_2$, $0 \le m \le n$. But then

$$D_\infty = \bigcap_{n=0}^{\infty} \operatorname{clos}\left(\bigcup_{j=n}^{\infty} D_j\right)$$

will satisfy the thesis of the lemma; a contradiction.

Recall that p denotes the canonical projection onto $\hat{M}$, and ϕ the suspension flow. Let $\xi = p(x,0)$ and $K_t = \{\eta \in M_t : V(\phi(\xi,t),\eta) < k\}$, $t \in \mathbf{R}$. For arbitrarily large $T > 0$, the connected component of K_T that contains $\phi(\xi,T)$, must have points η such that $\phi(\eta, t-T) \notin \operatorname{clos}(K_t)$ for some t, $0 \le t \le T$. Indeed, if this were not so, for sufficiently large T and every η in the mentioned component of K_T, $V(\phi(\xi, t+T), \phi(\eta,t)) \le k$, $-T \le t \le 0$; if $\varsigma = p(z,0)$, $z \in M$, is an ω-limit point of ξ we get that $V(\phi(\varsigma,t),\phi(\eta,t)) \le k$, for $t \le 0$ and η in some neighbourhood of ς. It follows easily that z would be an unstable point of f in contradiction with lemma 1.4.

Choose then $T > N$ and η, such that the above mentioned property holds. Let $a: [0,1] \to K_T$ be an arc, $a(0) = \phi(\xi,T)$, $a(1) = \eta$, and s^* the supremum of those $s \in [0,1]$ such that, for each u in $[0,s]$, $\phi(a(u), t-T) \in K_t$, $0 \le t \le T$, and $\phi(a(u), -T) \in A$; since $\ddot{V} > 0$, $\phi(a(s^*), t-T) \in K_T$ for $0 < t \le T$ and consequently, $\phi(a(s^*), -T) \in \partial A$. Thus, if $b = \phi(a([0,s^*]), -T)$, $f^n(b) \in B_{\delta_2}$ for $0 \le n \le N$, which is absurd.

For $x \in M$ and $0 < \delta < \alpha$, let $S_\delta(x)$, the δ-stable set of x, be

$$S_\delta(x) = \{y \in M : \operatorname{dist}(f^n(x), f^n(y)) \le \delta,\ n \ge 0\}.$$

Lemma 2.2. *Let $0 < \delta' < \delta$; there exists $\sigma > 0$ such that if $y \in S_\delta(x)$ and $\operatorname{dist}(x,y) < \sigma$, then $y \in S_{\delta'}(x)$.*

Proof. Take $k > 0$ such that for every $x, y \in M$, $V(x,y) < k$ implies $\operatorname{dist}(x,y) < \delta'$. Let $y \in S_\delta(x)$; if for some $t^* > 0$, $\dot{V}(\phi(x,t^*), \phi(y,t^*)) > 0$, the same is true for $t > t^*$ and therefore $V(\phi(x,t),\phi(y,t))$ would be increasing for $t > t^*$ which leads easily to a contradiction with expansivity. Thus $\dot{V}(\phi(x,t),\phi(y,t)) \le 0$ for $t \ge 0$ and consequently, for these t,

$$V(\phi(x,t),\phi(y,t)) \le k$$

provided that $V(x,y) \le k$. Choose σ such that $B_\sigma(x) \subset \{y : V(x,y) \le k\}$.

From now on we shall assume that α is so small that any subset of diameter α is contained in the domain of a coordinate map of M; also we shall consider these subsets of M as subspaces of the metric space $\mathbf{R}^2$ (or $\mathbf{S}^2$) and talk about coordinate axis, straight lines, etc.

Let $C_\delta(x)$ be the connected component of $S_\delta(x)$ such that $x \in C_\delta(x)$.

Lemma 2.3. *$C_\delta(x)$ is locally connected at x.*

Proof. Let $\sigma_o > 0$ be so small that for any $0 < \sigma < \sigma_0$ we have that

i) for $y \in B_\sigma(x)$ there is a connected set joining y to $\partial B_\sigma(x)$, contained in $\operatorname{clos}(B_\sigma(x)) \cap S_{\delta/2}(y)$ (Lemma 2.1), and

ii) if $y \in S_\delta(x)$ and $y \in B_\sigma(x)$ then $y \in S_{\delta/2}(x)$ (Lemma 2.2).

For such a σ and y, let D_x be the connected component of $\operatorname{clos}(B_\sigma(x)) \cap C_\delta(x)$ including x, and if y belongs to $C_\delta(x)$ let us call D_y the connected component of $\operatorname{clos}(B_\sigma(x)) \cap C_\delta(x)$ that contains y; on account of i), we may assert that both D_x and D_y meet the boundary of $B_\delta(x)$.

Arguing again by contradiction, we may assume that for some σ in the above conditions there are $y \in C_\delta(x)$ arbitrarily close to x, such that $D_x \cap D_y$ is empty. Let S' be the straight line segment joining x to such a y close to x. There is a first point z in S' that belongs to D_y and a point $u \in D_x \cap S'$ such that the segment $S \subset S'$ joining u to z meets D_y only in z and D_x only in u. Analogously, there is a small arc A of the circumference $\partial B_\sigma(x)$ that meets $D_x \cup S \cup D_y$ only in two points, one in D_x and the other one in D_y. As $D_x \cup S \cup D_y$ is connected, $D_x \cup S \cup D_y \cup A$ separates $\mathbf{R}^2$ (see [6], p. 506), but since neither D_x nor D_y separate $\mathbf{R}^2$ – for otherwise f would have a stable point – no union of three of the four D_x, D_y, S, A, separates $\mathbf{R}^2$, as may be shown applying theorem 7 of [6], p. 507. Therefore, the boundary of any component of the complement of $H = H_y = D_x \cup S \cup D_y \cup A$, must include the arc A; it follows easily that there are only two components of the complement of H, one of them contains the exterior of $B_\sigma(x)$, while the other one, say $G = G_y$, is bounded.

On account of the compactness of D_x it is easy to see that there exists $\gamma > 0$ with the following property: if $u, v \in D_x$, $\operatorname{dist}(u,v) \ge \sigma/8$, then $\operatorname{dist}(U_{\delta/2}(w), u)$
$> \gamma$ for any w such that $\operatorname{dist}(w,v) \le \gamma$. Here $U_{\delta/2}(w)$ denotes the $\delta/2$-unstable set of w; the δ-unstable set of a point x is defined as

$$U_\delta(x) = \{y : \operatorname{dist}(f^n(x), f^n(y)) \le \delta,\ n \le 0\}.$$

Take such a $\gamma < \sigma/8$ and choose $\rho > 0$ such that if $\operatorname{dist}(x,y) < \rho$, the corresponding component G_y is contained in the set of points whose distance to D_x is less than γ. The existence of such a ρ may be shown in the following way: each $z \in B_\sigma(x)$, $z \notin D_x$ may be joined to the exterior of $B_\sigma(x)$ by a compact arc a included in the complement of D_x; as any connected set that contains x and is contained in $S_\delta(x) \cap B_\sigma(x)$ must be included in D_x, if y is close enough to x, and $y \in C_\delta(x)$, $D_x \cap D_y = \varnothing$, then the corresponding H_y

does not meet the arc a. For, otherwise, we take a sequence of y_n satisfying the same conditions as the previous, $y_n \to x$, and show that the intersection of the sets $F_n = \mathrm{clos}(\cup_{j=n}^{\infty} D_{y_j})$, is a connected set, and therefore contained in D_x, that meets the arc a; a contradiction. Thus, there exists a $\rho_x > 0$, such that if $\mathrm{dist}(y, D_x) < \rho_x$ there is a neighbourhood of z disjoint from G_y; then a compactness argument on the set $\{z \in B_\sigma(x) \colon \mathrm{dist}(z, D_x) \geq \gamma\}$ permits to find such a ρ. Pick $y \in C_\delta(x)$, $D_x \cap D_y = \varnothing$, such that $\mathrm{dist}(G_y, D_x) < \gamma$, and denote G_y again by G.

Let P be a poligonal line of sides parallel to the coordinate axes, that joins an interior point of S to an interior point of A, and is included in G, except for its endpoints. Let Q be a straight line such that

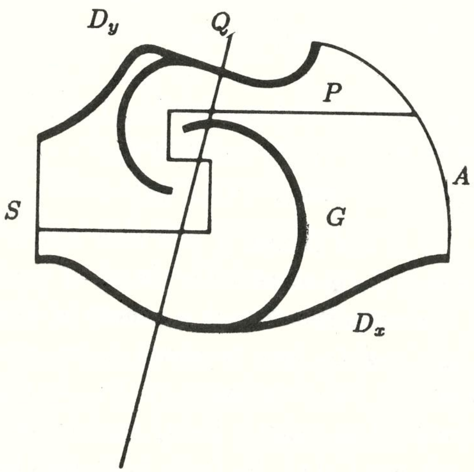
i) $\mathrm{dist}(Q, S \cup A) > \sigma/4$.

ii) $Q \cap P \neq \varnothing$ but $Q \cap P$ contains no vertex of P.

iii) Q separates S from A.

If $w \in Q \cap G$, $U_{\delta/2}(w)$ does not cut $S \cup A$ because there exists $v \in D_x$, $\mathrm{dist}(v, w) \leq \gamma$ so $\mathrm{dist}(v, x) \geq |\mathrm{dist}(x, w) - \mathrm{dist}(v, w)| \geq \sigma/4 - \gamma \geq \sigma/8$ and according to the above mentioned properties characterizing γ, this implies $\mathrm{dist}(U_{\delta/2}(w), x) > \gamma$; similarly $\mathrm{dist}(v, u) \geq \sigma/8$ for $\{u\} = D_x \cap A$.

Q cuts P in a finite set; the number of points in this set must be odd for otherwise, S and A would lie in the same component of the complement of Q, in contradiction with iii). On the other hand $Q \cap G$ consists of a union of disjoint open segments of Q, and one of them must contain an odd number of points of P. Assume that one of the endpoints of such a segment is in D_x.



Then we claim that the other one lies on D_y. Indeed, consider $H_1 = D_y \cup P \cup S_1 \cup A_1$ where $S_1 \subset S$, $A_1 \subset A$, are segments with one endpoint in D_y and the other one in P. Since D_x is connected and does not meet H_1, if both endpoints of the segment belonged to D_x they would belong to the same component of the complement of H_1; but this is impossible, since each time the segment meets P, it goes from one component of H_1 to the other one. Thus, we have found an open segment Q° contained in G, whose endpoints belong to D_x and D_y.

Let Q_1° be the subset of Q° that consist of those points whose $\delta/2$-unstable set does not meet D_x; Q_1° is open and also non-empty, since the endpoint of Q_1° that belongs to D_y has, because of expansivity, the property defining Q_1°. Similarly, $Q^\circ - Q_1^\circ$ consists of those points whose $\delta/2$-unstable set does not meet D_y, and is also open and non-empty. (Recall that, by previous arguments, these unstable sets do not meet $S \cup A$). Thus, we get a contradiction.

Assume now that $\delta < \alpha/3$ and let $y \in C_\delta(x)$; then $C_\delta(x) \subset C_{2\delta}(y)$. By the previous lemma, $C_{2\delta}(y)$ is locally connected at y, and therefore for any $\sigma > 0$ and any z close enough to $y, z \in C_\delta(x)$, there exists a compact connected set C joining y to z such that $C \subset B_\sigma(y) \cap C_{2\delta}(y)$. $C_\delta(x) \cup C \subset C_{2\delta}(y)$ and so, $C_\delta(x) \cup C$ can not separate $\mathbf{R}^2$. Consequently, by [6], p. 506, $C_\delta(x) \cap C$ must be connected; thus, $C_\delta(x) \cap C$ is a connected set joining y to z within $B_\sigma(y)$. We have then proved the following corollary.

Corollary 2.4. *For any $x \in M$, $C_\delta(x)$ is compact connected and locally connected.*

In the sequel, arc will mean a homeomorphic image of $[0,1]$.

Corollary 2.5. *For x in M and any two points $p, q \in C_\delta(x)$ there is an arc, contained in $C_\delta(x)$, joining p to q.*

Proof. See [6] Section 50.

3. Local product structure

Let σ, δ, $0 < \sigma < \delta < \alpha$, be so that for any $x \in M$ and $z \in B_\sigma(x)$, the intersections $S_{\delta/2}(z) \cap \partial B_\sigma(x)$ and $U_{\delta/2}(z) \cap \partial B_\sigma(x)$ are not empty, and such that if $z \in S_\delta(x)$, then $z \in S_{\delta/2}(x)$. Call $C = C(x, \sigma)$ the connected component of $C_\delta(x) \cap B_\sigma(x)$ that contains x; from the previous arguments it follows that C is locally connected and that any two points on C may be joined within C, by an arc.

Consider the family A of all arcs contained in C and having their origins at x and their endpoints on $\partial B_\sigma(x)$. When two arcs of A meet at a point different than x, they have in common an arc through x because C does not separate $\mathbf{R}^2$ and has empty interior. Among the arcs of A we define an equivalence relation according to which the arc $\hat{a}$ is equivalent to the arc $\hat{b}$ if $\hat{a} \cap \hat{b}$ strictly contains $\{x\}$.

Lemma 3.1. *There are only a finite number of equivalence classes.*

Proof. Let $\rho > 0$ be so that

$$\operatorname{dist}(U_\delta(x) \cap \partial B_\sigma(x),\ \ C \cap \partial B_\sigma(x)) > \rho,$$

and assume that for two non-equivalent arcs $\hat{a}$, $\hat{b} \in A$, their endpoints determine on $\partial B_\sigma(x)$ a compact arc $\hat{c}$ with diameter less than ρ. Then $U_\delta(x)$ does not cut $\hat{c}$. Let X be the union of the curve $\hat{a} \cup \hat{b} \cup \hat{c}$ with its interior and D the connected component containing x of $U_{\delta/2}(x) \cap X$. There is an open connected neighbourhood N of D in X, such that for every z in the closure of N, $U_{\delta/2}(z) \cap \hat{c}$ is void. By [6], p. 437, the boundary B of the connected component of $X - N$ that contains $\hat{c}$, is connected; $x \notin B$ and also, the intersections $B \cap \hat{a}$ and $B \cap \hat{b}$ are not empty. Since for $z \in B$, $U_{\delta/2}(z) \cap \partial B_\sigma(x) \neq \emptyset$ and $U_{\delta/2}(z) \cap \hat{c} = \emptyset$, $U_{\delta/2}(z)$ must cut either $\hat{a}$ or $\hat{b}$. Then the points of B may be classified according to whether the $\delta/2$-unstable sets through them meet $\hat{a}$ or $\hat{b}$; both classes are open and non-void, which is absurd. Thus, any set of endpoints of representatives of different classes must be finite.

Assume now that, at x, there are least two equivalence classes of arcs of A. Then it is easy to see that there are non-equivalent arcs $\hat{a}, \hat{b} \in A$ such that for some arc $\hat{c} \subset \partial B_\sigma(x)$, $\hat{a} \cup \hat{c} \cup \hat{b}$ is a Jordan curve and C only meets $\hat{c}$ at its endpoints. Call again X the union of $\hat{a} \cup \hat{c} \cup \hat{b}$ with its interior and D the connected component of $U_{\delta/2}(x) \cap X$ that contains $\{x\}$.

Lemma 3.2. *D separates X and therefore there is an arc in D joining x to a point of $\hat{c}$.*

Proof. If D does not meet $\hat{c}$ we may repeat the connectedness argument in the proof of the previous lemma to reach a contradiction. The last assertion follows from the results of the preceding section applied to f^{-1}.

Let $\hat{c}_1, \hat{c}_2$ be arcs contained in $\hat{c}$ so that $\hat{c}_1$ begins at the endpoint of $\hat{a}, \hat{c}_2$ ends at the endpoint of $\hat{b}$, and $D \cap (\hat{c}_1 \cup \hat{c}_2) = \emptyset$. Let N be an open connected neighbourhood of x in X such that for $y \in N$ the connected component of $S_{\delta/2}(y) \cap \hat{c}$ that contains $\{y\}$ is, in turn, included in $\hat{c}_1 \cup \hat{c}_2$. Moreover, we choose N and $\hat{c}_1, \hat{c}_2$ such that the $\delta/2$-unstable set through any point of N does not meet $\hat{c}$ in points that belong to $\hat{c}_1 \cup \hat{c}_2$. Let Q be the subset of N that consist of those y such that

i) There is an arc $\hat{s}(y)$ through y, $\hat{s}(y) \subset S_{\delta/2}(y)$ that intersects $\hat{c}_1$ and $\hat{c}_2$,

ii) There is an arc $\hat{u}(y)$ through y, $\hat{u}(y) \subset U_{\delta/2}(y)$, that meets $\hat{c}$ and $\partial B_\sigma(x) - \hat{c}$.

In the sequel, we shall say that $y \in M$ has a local product structure if there is a homeomorphism of $\mathbf{R}^2$ onto an open neighbourhood of y such that it maps horizontal (vertical) lines onto open subsets of local stable (resp. unstable) sets.

Lemma 3.3. *Q is open and non-void and every $y \in Q \cap \operatorname{int}(X)$ has a local product structure.*

Proof. Let us first show that $x \in Q$. In the presence of non-equivalent arcs $\hat{a}$ and $\hat{b}$, it is clear that we can find a set X', analogous to X, bordered by arcs $\hat{a}', \hat{b}', \hat{c}'$; $\hat{a}', \hat{b}' \subset S_{\delta/2}(x)$; $\hat{c}' \subset \partial B_\sigma(x)$, and such that c' and $\hat{c}$ have no interior point in common. By lemma 3.2, there exists $D' \subset X'$ defined in a similar way as our previous $D, D' \cap D = \{x\}$ and an arc contained in D' and joining x to an interior point of $\hat{c}'$. This proves that x satisfies ii); as i) is obviously true for x, we get $x \in Q$.

Now we show that Q is open in N. Let $y \in Q$; for z close to $y, z \in \hat{u}(y)$, the arguments of the previous paragraph may be applied to get that the connected component of $S_{\delta/2}(z) \cap B_\sigma(x)$ that contains $\{z\}$ meets $\hat{c}_1$ and $\hat{c}_2$. Similarly, for t close to y but on $\hat{s}(y)$, the corresponding connected component has to cut $\hat{c}$ and $\partial B_\sigma(x) - \hat{c}$. The function that sends (z, t) to the unique intersection point of $S_{\delta/2}(z)$ and $U_{\delta/2}(t)$ is continuous, injective and open, by invariance of domain. This proves the lemma.

Corollary 3.4. *Let $x \in M$ be such that there are at least two non-equivalent arcs in A. Then there is a neighbourhood of x such that each y in that neighbourhood $y \neq x$, has a local product structure.*

Proof. To construct such a neighbourhood it is enough to remark that if $a \in A$ there is a non trivial arc $\hat{e}$, beginning at x, such that if $\hat{a}'$ is equivalent to $\hat{a}$, then $\hat{e} \subset \hat{a}'$ (see the proof of Lemma 3.1), and to replace stable arcs by unstable ones to include the points in the border of sectors such as our previous X.

Lemma 3.5. *For each $x \in M$ there is $\sigma > 0$ such that the family A has at least two non-equivalent arcs.*

Proof. Assume that for some x we have that for every $\sigma > 0$ the family A defined by x and σ has only one class of arcs. Pick a small σ; let the arc $\hat{a}$ with origin x and endpoint $z \in \partial B_\sigma(x)$ be a representative of that class. If C, the connected component of $S_\delta(x) \cap B_\sigma(x)$ that contains $\{x\}$, also contains points other than those in $\hat{a}$, we join them to x, withn C, by all possible arcs. If all these arcs contain $\hat{a}$, it is easy to see that for some smaller σ, C would consist of only one arc joining x to $\partial B_\sigma(x)$. Assume then that there is a point $v \in C$, $v \notin \hat{a}$, that may be joined to x, within C, by an arc which does not contain $\hat{a}$. Thus, there is a point $u \in \hat{a}$, $u \neq x$, and an arc $\hat{b} \subset C$ with origin u and endpoint v, whose intersection with $\hat{a}$ is $\{u\}$. Let J be a Jordan curve through z and v such that $\hat{a}$ and $\hat{b}$ lie in its interior except for their endpoints. Let $w \in \hat{a}$, $w \neq x$ be the closest point to x such that there is an arc $\hat{c} \subset C$ with origin w and endpoint on J, $\hat{c} \cap \hat{a} = \{w\}$ (lemma 3.1). Consequently, the arc contained in $\hat{a}$, with origin x and endpoint w, belongs (except for w) to the interior of one of the stable sectors, say X, defined as previously, with w replacing x and J instead of $\partial B_\sigma(x)$. But on account of the local product structure on a neighbourhood of w in X, this implies that the stable set of w meets twice some unstable set, which is absurd.

Thus for some $\sigma > 0$, C consists of an arc $\hat{a}$ interior to $B_\sigma(x)$ except for its endpoint. All interior points of $\hat{a}$ have a local product structure and therefore their local unstable sets are arcs transversal to $\hat{a}$. A connectedness argument on the boundary of a small neighbourhood of $\hat{a}$ (see [6], p. 437) permits to find a stable arc that has to cut $\partial B_\sigma(x)$ up and down $\hat{a}$ (but close to $\hat{a}$): if not, we break the connectedness of the boundary of the mentioned small neighbourhood of $\hat{a}$. These arcs meet twice some of those local unstable arcs.

Proposition 3.6. *Except for a finite number of points, that we shall call singular, every $x \in M$ has a local product structure. Any singular point y is periodic and its local stable (unstable) set consist of the union of r arcs that meet only at y; $r \geq 3$. The stable (unstable) arcs separate unstable (resp. stable) sectors.*

Proof. Lemma 3.5 and corollary 3.4 imply that singular points can not accumulate; thus, there are only a finite number of them. Since the set of singular points is f-invariant, all of them are periodic. The rest of the assertions are also easy consequences of our previous results.

4. The sphere

A rectangle in M will be the image of $[0,1]\times[0,1]$ through a homeomorphism that maps horizontal (vertical) segments into stable (resp. unstable) arcs; we assume that a rectangle has a diameter less than α and that it may contain at most one singular point, and only as one of its vertices.

By standard procedures we may construct a finite family R of rectangles R_i, whose union is M, and such that, if $i \neq j$, $R_i \cap R_j = \partial R_i \cap \partial R_j$, and that this intersection consists of an arc included in one of the sides of each rectangle, when it is non-void. Take a point x_1, on an unstable side of, say, R_1, (it may be a

vertex), such that it belongs to no stable leaf through any singular point. In the unstable side of R_1 opposite to that of x_1 and on the stable arc of R_1 through x_1, we take x_2; since x_2 belongs to the unstable side of another rectangle, say R_2, we may continue with the same procedure. We find, in this way, a finite collection of rectangles and points, that we shall denote by R_i, x_i, $i = 1, \ldots, n$, so that

i) $R_i \neq R_j$ if $i \neq j$, $i,j = 1, \ldots, n-1$, and $R_1 = R_n$.

ii) For $i = 2, \ldots, n$, x_i belongs to an unstable side of R_{i-1} and of R_i.

In particular, x_n belongs either to the unstable side of R_1 through x_1 or to the one where x_2 lies. In the first case we define x_{n+1} continuing, from x_n, the previous procedure. Then, in this case, the union of the stable arcs $x_2x_3, \ldots, x_{n-1}x_n$ with the unstable arc $x_{n+1}x_2$ is a Jordan curve, as well as, in the other case, the union of the stable arcs $x_2x_3, \ldots, x_{n-1}x_n$ with the unstable x_nx_2.

Let M be homeomorphic to S^2 and let C be a connected component of the complement, say, the interior, of one of the Jordan curves defined previously.

For any $x \in M$ there is a sequence x_k, $k = 0, 1, \ldots$, so that $x_0 = x$, x_kx_{k+1} is a stable arc contained in a rectangle of R and $x_kx_{k+1} \cap x_{k'}x_{k'+1}$ consists of at most one point if $k \neq k', k, k' = 0, 1, \ldots$. Moreover, for $k = 0, 1, \ldots$, x_k and x_{k+1} lie on opposite sides of some rectangle. The existence of such a sequence (we shall call the union of the arcs x_kx_{k+1} an infinite stable prolongation of x_0x_1; an H stable prolongation, if $k = 0, 1, \ldots, H$), is an easy consequence of the non-existence of stable closed curves.

Lemma 4.1. *C contains some infinite stable prolongation*

Proof. Assume that C contains some singular point. If the thesis of the lemma were not true, we could find a Jordan curve J, $J = a_1 \cup a_2 \cup b$, where a_1, a_2 are arcs of the stable set of a singular point p, $a_1 \cap a_2 = \{p\}$, and b is an arc included in $x_{n+1}x_2$ (or in x_nx_2). Moreover, we may choose J so that no singular point belongs to the interior of J (contained in C). Then, the continuous function that sends $x \in b$ to the first point $y \in b$ where the stable prolongation of x entering C meets b again, would have a fixed point. Since this is absurd, and as the case when C contains no singular point may be handled in the same way, this completes the proof.

Theorem 4.2. *There are no expansive homeomorphisms of* S^2.

Proof. Let f be an expansive homeomorphism of S^2 and let C be as above. The previous lemma permits to construct in the same way as the border of C, a Jordan curve whose interior, C_1, is properly included in C and contains an infinite stable prolongation. Consider a maximal family of sets like $\mathrm{clos}(C)$, $\mathrm{clos}(C_1)$, ordered by inclusion. It is easy to see that the intersection of such a family contains an infinite stable prolongation, and, then, it contains properly a set that belongs to the family: a contradiction.

5. The torus

Let $\overline{f}$ be a lifting of f to the universal covering $\overline{M} \simeq \mathbf{R}^2$ or $\mathbf{S}^2$ of M and let $\overline{s}$, $\overline{u}$, be, respectively, the corresponding liftings of a stable and an unstable leaf of f. Clearly $\overline{s}(\overline{u})$ is a stable (unstable) leaf of $\overline{f}$. If $\overline{s}$ and $\overline{u}$ meet at two points we find a Jordan curve whose boundary is the union of an arc of $\overline{s}$ with an arc of $\overline{u}$ and therefore, by previous arguments, it contains an infinite stable (or unstable) semi-leaf. Since the reasoning of lemma 3.2 applies essentially in the same way, this is absurd; therefore we may state:

Corollary 5.1. *$\overline{s} \cap \overline{u}$ is either empty or consists of one point.*

Now we assume that $M = \mathbf{T}^2$ and that f is homotopic to an Anosov diffeomorphism g; by [3], f is semi-conjugate to g. Let then $h: \mathbf{T}^2 \to \mathbf{T}^2$ be a continuous surjective map, $h \circ f = g \circ h$, and let $\overline{f}, \overline{g}, \overline{h}$ be liftings of f, g, h such that $\overline{h}: \mathbf{R}^2 \to \mathbf{R}^2$ is a proper map that satisfies $\overline{h} \circ \overline{f} = \overline{g} \circ \overline{h}$ (see [3], p. 62-64).

Let R be a family of rectangles of $\mathbf{T}^2$ with the properties of the previous one, and moreover, such that if an unstable (stable) arc β is contained in the union of two rectangles of R, then, $\mathrm{diam}(f^n(\beta))$ is less than the expansivity constant α for any $n \leq 0$ (resp. $n \geq 0$). Let $\overline{R}$ be the lifting of R and assume that $\overline{y}$ belongs to the unstable leaf $\overline{u}$ through $\overline{x}$, $\overline{x} \neq \overline{y}$. Let β be the arc of u with endpoints $\overline{x}, \overline{y}$; then $\mathrm{diam}(\overline{f}^n(\beta)) \to \infty$, for otherwise there would exist a positive integer K and a subsequence n_i, $i = 1, 2 \ldots$, such that $\overline{f}^m(\beta)$ would be contained, for $m = n_i$, in the union of K rectangles of $\overline{R}$. Choose a partition of β as the union of more than K sub-arcs; we have, on account of the properties of R with respect to α, that there exists N such that for $n \geq N$, the $\overline{f}^n$ image of any one of the arcs of the partition contains an entire unstable arc of some rectangle of $\overline{R}$ (if not there would be an arc say γ, such that for $n \geq 0$ $\mathrm{diam} f^n(\gamma) \leq \alpha$, which is absurd). Since by previous arguments, $\overline{f}^n(\overline{u})$ can not meet twice a rectangle of $\overline{R}$ we get that for $n \geq N$, $\overline{f}^n(\beta)$ can not be included in the union of K rectangles.

Now we prove the following lemma which admits a similar version for points on a stable arc and $n \to -\infty$.

Lemma 5.2. $\lim\limits_{n\to\infty} \mathrm{dist}(\overline{f}^n(\overline{x}), \overline{f}^n(\overline{y})) = \infty.$

Proof. If this were not true we could find another sequence $m_i \to \infty$ of positive

integers m_i, a compact set $L \subset \overline{M}$, and, for each $m = m_i$, points $\overline{x}_m, \overline{y}_m \in L$, such that

i) $\overline{x}_m = \overline{f}^m(\overline{x}) + e_m$, $\overline{y}_m = \overline{f}^m(\overline{y}) + e_m$, where $e_m \in \mathbb{Z} \times \mathbb{Z}$, and

ii) the diameter of the unstable arcs with endpoints $\overline{x}_m, \overline{y}_m$ tends to infinite.

In these conditions it is easy, on account of ii), to find points $\overline{z}_m, \overline{w}_m$ in the straight line segment joining $\overline{x}_m$ to $\overline{y}_m$, such that the stable leaf through $\overline{z}_m$ and the unstable leaf through $\overline{w}_m$ meet at points whose distance to $\overline{z}_m$ or $\overline{w}_m$ tends to infinite with m. But on account of the properties of Anosov diffeomorphisms i.e., for $\overline{x}, \overline{y} \in \mathbf{R}^2$ the stable leaf of $\overline{x}$ and the unstable leaf of $\overline{y}$ meet at exactly one point (see [3], p. 89) and the fact that $\overline{h}$ is a proper map, this is absurd as it may be shown easily by taking a subsequence of m_i such that the corresponding z_m and w_m converge.

It follows from the preceding lemma that if $\overline{y}$ lies in the unstable set of $\overline{x}$, whether this set contains singular points or not, $\overline{h}(\overline{x}) \neq \overline{h}(\overline{y})$ provided $\overline{x} \neq \overline{y}$. For otherwise we would have for $n \geq 0$,

$$\overline{h}(\overline{f}^n(\overline{x})) = \overline{g}^n(\overline{h}(\overline{x})) = \overline{g}^n(\overline{h}(\overline{y})) = \overline{h}^n(\overline{f}(\overline{y}))$$

and then, we could find $e_n \in \mathbb{Z} \times \mathbb{Z}$ such that $\overline{f}^n(\overline{x}) + e_n$, $\overline{f}^n(\overline{y}) + e_n$ lie in the $\overline{h}$ pre-image of the fundamental square in $\mathbf{R}^2$. But since $\mathrm{dist}(\overline{f}^n(\overline{x}), \overline{f}^n(\overline{y})) \to \infty$, this is absurd.

Consequently f has no singular points, and this, in turn implies that $\overline{h}$ is a covering map since if $\overline{x}, \overline{z}$ have the same $\overline{h}$ image and $\overline{z}$ lies in the neighbourhood of $\overline{x}$ with a local product structure we could find $\overline{y}$ in the unstable leaf through $\overline{x}$ and such that $\overline{h}(\overline{x}) = \overline{h}(\overline{y})$ and

$$\sup_{n\geq 0} \mathrm{dist}(\overline{f}^n(x), \overline{f}^n(y)) < \infty,$$

in contradiction with lemma 5.2. Thus h is a homeomorphism. We have proved then the following theorem.

Theorem 5.3. *If f is expansive and homotopic to an Anosov diffeomorphism g of $\mathbf{T}^2$, then f is conjugate to g.*

Let $f: \mathbf{T}^2 \to \mathbf{T}^2$ be expansive and $\overline{f}: \mathbf{R}^2 \to \mathbf{R}^2$ a lifting of f. Let A be the matrix of integer elements and determinant ± 1 such that $\overline{f}(\overline{x} + \nu) = \overline{f}(\overline{x}) + A\nu$ for every $\overline{x} \in \mathbf{R}^2$, $\nu \in \mathbb{Z} \times \mathbb{Z}$. Let $H > 0$ be such that $\left\|\overline{f}(\overline{x})\right\| \leq H$ for every $\overline{x}$ in the fundamental square, and let $\overline{x} \in \mathbf{R}^2$, $\overline{x} = \overline{x}_0 + \nu_0$, where $\overline{x}_0$ belongs to the fundamental square and $\nu_0 \in \mathbb{Z} \times \mathbb{Z}$. Then $\overline{f}(\overline{x}) = \overline{f}(\overline{x}_0) + A\nu_0$ and $\overline{f}(\overline{x}_0) = \overline{x}_1 + \nu_1$ where $\overline{x}_1$ belongs also to the fundamental square and $\nu_1 \in \mathbb{Z} \times \mathbb{Z}$, $\|\nu_1\| \leq H + 1$. If for $i = 0, 1, \ldots$, we define $\overline{x}_{i+1}$ by $\overline{f}(\overline{x}_i) = \overline{x}_{i+1} + \overline{\nu}_{i+1}$ where x_{i+1} belongs to the fundamental square $\nu_{i+1} \in \mathbb{Z} \times \mathbb{Z}$, $\|v_{i+1}\| \leq H + 1$, we have that

$$\overline{f}^{i}(\overline{x}) = \overline{x}_i + \nu_i + A\nu_{i-1} + \cdots + A^i\nu_0,$$

and if for some $s > 0$, $\left\| A^i \right\| \leq si$, we get that, for some $C, D > 0$,

$$\left\| \overline{f}^{i}(x) \right\| \leq C + \frac{1}{2}Di(i+1).$$

Let R be a finite family of rectangles, as before. However we choose this time the rectangles so small that no unstable arc included in the union of 4 rectangles of $\overline{R}$ can have a $\overline{f}^{n}$-image, $n \leq 0$, with diameter greater than σ, for some σ, $0 < \sigma < \alpha$. Let $m > 0$ be such that the $\overline{f}^{m}$-image of any unstable segment of a rectangle (i.e. an unstable arc, contained in the rectangle, and joining opposite stable sides of it) in $\overline{R}$ meets at least 4 rectangles; consequently, this image contains unstable segments of 2 rectangles. Since by previous arguments no unstable leaf may contain points on different unstable arcs of a $\overline{R}$ rectangle, $\overline{f}^{km}(\hat{b})$ contains unstable segments of at least 2^k rectangles, provided $\hat{b}$ is itself, an unstable segment. On the other hand, if for $s > 0$, $n = 1, \ldots$ we had $\|A^n\| \leq sn$, the number of rectangles met by $\overline{f}^{km}(\hat{b})$ would be, according to the previous inequalities, less than Bk^4 for some $B > 0$. Thus, A must be hyperbolic (for otherwise $\left\| A^i \right\| \leq si$ for some $s > 0$ and $i = 1, 2 \ldots$).

Lemma 5.4. *Let $f\colon \mathbf{T}^2 \to \mathbf{T}^2$ be an expansive homeomorphism; then f is homotopic to an Anosov diffeomorphism.*

Proof. The linear map $A\colon \mathbf{R}^2 \to \mathbf{R}^2$ considered above, is the lifting of a linear Anosov isomorphism g of $\mathbf{T}^2$; f is homotopic to g.

Theorem 5.5. *Let $f\colon \mathbf{T}^2 \to \mathbf{T}^2$ be an expansive homeomorphism; then f is conjugate to an Anosov diffeomorphism.*

6. Surfaces of genus ≥ 2

Let M be a surface of genus greater then 1, and $f\colon M \to M$ an expansive homeomorphism. Let $g\colon M \to M$ be a pseudo-Anosov map isotopic to f, $\overline{M}$ the universal covering of M, $\overline{f}$, $\overline{g}$ suitable liftings of f and g to $\overline{M}$, and $\overline{D}_s$, $\overline{D}_u$ equivariant pseudo-metrics $\overline{D}_s, \overline{D}_u\colon \overline{M} \times \overline{M} \to \mathbf{R}$ such that $\overline{D}_s(\overline{x},\overline{y})$ ($\overline{D}_u(\overline{x},\overline{y})$) is the infimum of the length in terms of the g-measure transverse to the unstable (resp. stable) foliation, of the arcs joining $\overline{x}$ to $\overline{y}$ [11, 2]. There exists $\lambda > 1$ such that

$$\overline{D}_s(\overline{g}^{-1}(\overline{x}), \overline{g}^{-1}(\overline{y})) = \lambda \overline{D}_s(\overline{x},\overline{y}), \overline{D}_u(\overline{g}(\overline{x}), \overline{g}(\overline{y})) = \lambda \overline{D}_u(\overline{x},\overline{y});$$

furthermore, $\overline{D} = \overline{D}_s + \overline{D}_u$ is an equivariant metric on $\overline{M}$.

Lemma 6.1. *There exists $P > 0$ such that for any $\overline{x}$ which does not belong to stable or unstable sets of the $\overline{f}$ singular points, there exists $\overline{y} \in \overline{M}$ such that $\overline{D}(\overline{f}^{n}(x), \overline{g}^{n}(y)) \leq P$, $n \geq 0$.*

Proof. For $P > 0$ and $\xi \in \overline{M}$, let $B_P(\xi)$, denote the $\overline{D}$ balls of radius P centered at ξ. Let $W_u^{\overline{f}}(\overline{x})$ be a closed arc of the unstable manifold through $\overline{x}, \overline{x} \in W_u^{\overline{f}}(\overline{x})$, that meets the boundary of $B_P(\overline{x})$ only at both endpoints. Assume that for no $\overline{y} \in W_u^{\overline{f}}(\overline{x})$, $\overline{D}(\overline{g}^{n}(y), \overline{f}^{n}(\overline{x})) \leq P$, $n \geq 0$. Then there exists $N > 0$, such that for every $\overline{y} \in W_u^{\overline{f}}(\overline{x})$ it is possible to find $n(\overline{y})$, $0 \leq n(\overline{y}) < N$, such that

$$\overline{D}(\overline{g}^{n(y)}(\overline{y}), \overline{f}^{n(y)}(\overline{x})) > P.$$

Let $\hat{M}, \phi$ be the suspension of $(\overline{M}, \overline{f})$ and let $\rho > 0$ be such that if $\mathrm{d}_\rho(\overline{x})$ denotes the usual open disk of $\overline{M}$, with radius ρ and centre $\overline{x}$, we have that $\phi_{n+t}(\mathrm{d}_\rho(\overline{x}))$ contains $\phi_t(B_P(\overline{f}^{n}(\overline{x}))$ for $0 \leq n \leq N$, $0 \leq t \leq 1$. Now we consider arcs $w_s(\phi(\overline{x},t))$ through $\phi(\overline{x},t)$, of the stable manifold of $\phi(\overline{x},t)$, such that their endpoints lie outside of $\phi_t(\mathrm{d}_\rho(\overline{x}))$, $0 \leq t \leq N$, and moreover, such that these arcs depend continuously on t. Also, let $\overline{U}(\overline{x})$ be a small closed arc of the $\overline{f}$ unstable manifold through $\overline{x}$, $\overline{x} \in \mathrm{int}(\overline{U}(\overline{x}))$, such that $\phi_t(\overline{U}(\overline{x})) \subset \phi_t(\mathrm{d}_\rho(\overline{x}))$, $0 \leq t \leq N$; $\phi_n(\overline{U}(\overline{x})) \subset B_P(\overline{f}^{n}(\overline{x}))$, $0 \leq n \leq N$. Let $Q(t)$, $0 \leq t \leq N$ be the ϕ_t image of one of the points of $\overline{U}(\overline{x})$; for points

$$\overline{z} \in \phi_t(\mathrm{d}_\rho(\overline{x})),\ \overline{z} \notin w_s(\phi(\overline{x},t)),\ \ 0 \leq t \leq N,$$

we consider the modulo 2 intersection number $i(\overline{z})$ of the arc $w_s(\phi(x,t))$ with an arc joining $\overline{z}$ to $Q(t)$ within $\phi_t(\mathrm{d}_\rho(\overline{x}))$. Let $a > 0$; if for $0 \leq t \leq a$, $\phi(\overline{z},t) \notin w_s(\phi(\overline{f}^{n}(\overline{x}),t))$ and $\phi(\overline{z},t) \in \phi_{n+t}(\mathrm{d}_\rho(\overline{x}))$, then clearly, $i(\phi(\overline{z},t))$ is constant on $[0,a]$. Thus, if $0 \leq n \leq N$, $\overline{D}(\overline{z}, \overline{f}^{n}(\overline{x})) \leq P$ and $\overline{z}$ does not belong to the whole stable manifold through $\overline{f}^{n}(\overline{x})$, $i(\overline{f}(\overline{z})) = i(\overline{z})$.

Let

$$R > \max\left\{ \sup_{x \in \overline{M}} \overline{D}(\overline{f}(\overline{x}), \overline{g}(\overline{x})),\ \ \sup_{x \in \overline{M}} \overline{D}(\overline{f}^{-1}(\overline{x}), \overline{g}^{-1}(\overline{x})) \right\},$$

$$H > 3R(\lambda - 1)^{-1} \quad \text{and} \quad P > (\lambda H + 2R)(\lambda - 1)^{-1}.$$

Assume that for $\overline{y}$ in $W_u(\overline{x})$ there is m, $0 \leq m \leq N$, such that

$$\overline{D}(\overline{g}^{n}(\overline{y}), \overline{f}^{n}(\overline{x})) \leq P$$

for $0 \leq n < m$ and $\overline{D}(\overline{g}^{m}(\overline{y}), \overline{f}^{m}(\overline{x})) \geq P$. Let $\overline{z}$ stand for $\overline{g}^{m}(y)$; if

$$\overline{D}_s(\overline{z}, \overline{f}^{m}(\overline{x})) > P - H$$

we would have, since $\overline{D}_s$ is a pseudometric, that

$$\begin{aligned}\overline{D}_s(\overline{g}^{-1}(\overline{z}),\overline{f}^{m-1}(\overline{x})) &\geq \overline{D}_s(\overline{g}^{-1}(\overline{z}),\overline{g}^{-1}(\overline{f}^{m}(\overline{x}))) - R\\ &= \lambda\overline{D}_s(\overline{z},\overline{f}^{m}(\overline{x})) - R\\ &\geq \lambda(P-H) - R\\ &> P\end{aligned}$$

which is absurd. Therefore $\overline{D}_u(\overline{z},\overline{f}^{m}(\overline{x})) \geq H$ and then

$$\begin{aligned}\overline{D}_u(\overline{f}(\overline{z}),\overline{f}^{m+1}(\overline{x})) &\geq \overline{D}_u(\overline{g}(\overline{z}),\overline{g}(\overline{f}^{m}(\overline{x})) - 2R\\ &\geq \lambda H - 2R \geq H + R\end{aligned}$$

Thus, $\overline{D}(\overline{f}^{k}(\overline{z}),\overline{f}^{m+k}(\overline{x})) \geq H$ for $k \geq 0$; it follows that $\overline{z}$ does not belong to the $\overline{f}$-stable manifold through $\overline{f}^{m}(\overline{x})$ and that, in case $\overline{D}(\overline{z},\overline{f}^{m}(\overline{x})) = P$, $i(\overline{z}) = i(\overline{f}(\overline{z}))$. On the other hand, if $\overline{D}(\overline{f}(\overline{z}),\overline{v}) \leq R$, we have that

$$\overline{D}_u(\overline{f}^{m+1}(\overline{x}),\overline{v}) + \overline{D}_u(\overline{v},\overline{f}(\overline{z})) \geq \overline{D}_u(\overline{f}^{m+1}(\overline{x}),f(\overline{z}))$$

which implies $\overline{D}_u(\overline{f}^{m+1}(\overline{x}),\overline{v}) \geq H + R - R = H$; therefore $\overline{v}$ does not belong to the entire $\overline{f}$-stable manifold of $\overline{f}^{m+1}(\overline{x})$ and $i(\overline{v}) = i(\overline{f}(\overline{z}))$. Since $\overline{D}(\overline{g}(\overline{z}),\overline{f}(\overline{z})) \leq R$ we get

$$i(\overline{g}(\overline{z})) = i(\overline{f}(\overline{z})) = i(\overline{z}).$$

Let $\overline{y} \in W_u^{\overline{f}}(x)$ and let $n(\overline{y})$, $0 < y < n(\overline{y}) < N$, be the first positive integer such that $\overline{D}(\overline{g}^{n(\overline{y})}(\overline{y}),\ \overline{f}^{n(\overline{y})}(\overline{x})) > P$. Since the previous arguments imply that $\overline{g}^{n(\overline{y})}(\overline{y})$ does not belong to the $\overline{f}$-stable manifold of $\overline{f}^{n(\overline{y})}(\overline{x})$ we may define $j(\overline{y})$ as the intersection number $i(g^{n(\overline{y})}(\overline{y}))$; $j(\overline{y}) = 0$ or 1. We want to show now that the sets

$$j_0 = \left\{\ \overline{y} \in W_u^{\overline{f}}(\overline{x}):\ j(\overline{y}) = 0\right\},$$

$$j_1 = \left\{\ \overline{y} \in W_u^{\overline{f}}(x):\ j(\overline{y}) = 1\right\}$$

are both open. If $\overline{y}$ is such that $\overline{D}(\overline{g}^{n}(\overline{y}),\overline{f}^{n}(\overline{x})) < P$ for each n, $0 \leq n \leq n(y)$, or if $\overline{D}(\overline{g}^{n}(\overline{y}),\overline{f}^{n}(\overline{x})) < P$ for $0 \leq n < n(y) - 1$ and

$$\overline{D}(\overline{g}^{n(\overline{y})-1}(\overline{y}),\overline{f}^{n(\overline{y})-1}(\overline{x})) = P$$

we know that $\overline{y}$ has a neighbourhood whose points $\overline{z}$ satisfy $j(\overline{z}) = j(\overline{y})$. So, let us try to choose P in order to have that if $\overline{D}(\overline{g}^{n}(\overline{y}).\overline{f}^{n}(\overline{x})) = P$ for some n, $0 \leq n \leq n(y)$, then $n = n(y) - 1$.

Consider, for $a, b \in \overline{M}$,

$$\begin{aligned}\Delta &= \overline{D}(\overline{g}(a),\overline{f}(b)) - 2\overline{D}(a,b) + \overline{D}(\overline{g}^{-1}(a),\overline{f}^{-1}(b))\\ &\geq \overline{D}(\overline{g}(a),\overline{g}(b)) - R - 2\overline{D}(a,b) + \overline{D}(\overline{g}^{-1}(a),\overline{g}^{-1}(b)) - R\\ &\geq (\lambda + \lambda^{-1} - 2)\overline{D}(a,b) - 2R.\end{aligned}$$

Let $r = \lambda + \lambda^{-1} - 2$ and choose P to satisfy also $P > 2r^{-1}R$; then if $\overline{D}(a,b) = P$, $\Delta > 0$. Thus if $\overline{D}(a,b) = P$ and $\overline{D}(\overline{g}^{-1}(a),\overline{f}^{-1}(b)) \leq P$ we get that $\overline{D}(\overline{g}(a),\overline{f}(b)) > P$. Hence, with this choice of P, we have that if for some n, $0 \leq n < n(y)$, $\overline{D}(\overline{g}^{n}(\overline{y}),\overline{f}^{n}(\overline{x})) = P$, then $n = n(y) - 1$, proving that j_0 and j_1 are open.

Let now e stand for an endpoint of $W_u^{\overline{f}}(x)$; if $\overline{D}_s(\overline{x},e) \geq H$, we get that $\overline{D}_s(\overline{f}^{-1}(x),\overline{f}^{-1}(e)) \geq \lambda H - 2R \geq H$, and that $\overline{D}_s(\overline{f}^{-k}(\overline{x}),\overline{f}^{-k}(e)) \geq H$, $k \geq 0$, which is absurd. Thus $\overline{D}_u(\overline{x},e) \geq P - H$ and therefore

$$\overline{D}(\overline{f}(\overline{x}),\overline{g}(e)) \geq \overline{D}_u(\overline{f}(\overline{x}),\overline{g}(e)) \geq \lambda(P-H) - R > P.$$

Hence it follows from the previous arguments that j_0 and j_1 are non-void; this completes the proof.

Lemma 6.2. *For each $\overline{x} \in \overline{M}$ there exists a unique $\overline{y} \in \overline{M}$ such that $\overline{D}(\overline{f}^{n}(\overline{x}),\overline{g}^{n}(\overline{y})) < P$, $n \in \mathbf{Z}$.*

Proof. From the previous lemma it follows easily that for each $\overline{x} \in \overline{M}$ which is not on the stable or unstable sets of the $\overline{f}$-singular points, there exist $\overline{y} \in \overline{M}$ such that $\overline{D}(\overline{f}^{n}(\overline{x}),\overline{g}^{n}(\overline{y})) \leq P$, $n \in \mathbf{Z}$. In fact, let $\overline{y}_n$ have, with respect to $\overline{f}^{-n}(\overline{x})$, $n \leq 0$, the property stated in that lemma; take y to be a limit point of $\overline{g}^{n}(y_n)$. Since those $\overline{x}$ constitute a dense subset of $\overline{M}$, we may complete the proof of the lemma using the expansivity properties of the liftings of pseudo Anosov maps i.e., that $\overline{D}(\overline{g}^{n}(\overline{x}),\overline{g}^{n}(\overline{y})) \leq k$, $n \in \mathbf{Z}$, implies $\overline{x} = \overline{y}$ [5].

Let $\overline{h}:\overline{x} \to \overline{y}$. then it follows from [5] (see, Theorems 1,2 and Lemma 2.2) that $\overline{h}$ is a surjective semiconjugacy between $\overline{f}$ and $\overline{g}$; furthermore $\overline{h}$ is a proper map. Let $\overline{R}$ be a family of rectangles of $\overline{f}$ as in the previous section. As in that section, we may show that $\overline{h}$ is injective on each rectangle of $\overline{R}$, on account of the fact that also for pseudo Anosov maps, if $\overline{x}_n \to \overline{x}$, $\overline{y}_n \to \overline{y}$ and the stable sets of $\overline{x}_n$ meets the unstable set of $\overline{y}_n$, $n \geq 0$, then the unstable set of $\overline{y}$, meets the stable set of $\overline{x}$. Since, by the same arguments, two prongs of a singular point can not go to the same prong of the image point, we conclude that every point $\overline{x}$ has a neighbourhood where $\overline{h}$ is injective; thus, $\overline{h}$ is a covering map. We have then proved the following theorem.

Theorem 6.3. *If $\overline{f}$ is expansive and homotopic to a pseudo Anosov $\overline{g}$, then $\overline{f}$ is conjugate to $\overline{g}$.*

Lemma 6.4. *An expansive homeomorphism can not preserve any non trivial isotopy class.*

Proof. Assume that for some expansive homeomorphism the assertion of the lemma is not true. Let f be an iterate of this homeomorphism such that there is a homeomorphic image of S^1, say σ, through a fixed (it is not difficult to show the existence of a periodic non-singular point of an expansive homeomorphism) non singular point x of f, whose non trivial isotopy class is preserved under f. We may construct, as in the proof of proposition II.6, exposé 5 of [2], a simple curve γ through x, isotopic to σ that consists of unstable arcs plus arcs transversal to both foliations plus singular points and is transversal to the stable foliation. Also we may construct another simple closed curve α through x, isotopic to σ, that consist of a finite union of stable and unstable arcs (not containing singular points). Clearly $f^n(\gamma)$ and $f^{-m}(\alpha)$ are isotopic, $n, m \geq 0$. Lemma 2.4 and 2.5 of [1] imply, for large n and m, the existence of a very large unstable arc Ω in $f^n(\gamma)$ that is continuously mapped, through the stable leaves starting at Ω, to a very small unstable arc $\Omega' \subset f^{-m}(\alpha)$. Indeed, this is a consequence of the existence of a disk (or a ring) bounded by an arc in $f^n(\gamma)$ and an arc in $f^{-m}(\alpha)$, our previous results concerning infinite prolongations of stable leaves contained in disks, together with the finiteness of the total number of prongs and the fact that the number of unstable arcs of $f^{-m}(\alpha)$, $m = 1, 2 \ldots$, is fixed.

Suppose now that we have a family of rectangles as the previous ones. Since, as it is easy to show, there is a positive integer H such that all the rectangles met by all possible infinite stable prolongations of the arcs joining opposite sides of any rectangle are also met after H stable prolongations (see section 4) of the same arcs, we may construct, starting from Ω, an unstable arc mapped continuously, following the stable leaves, on a part of it. But this implies the existence of a closed stable leaf, which is absurd.

Theorem 6.5. *Every expansive homeomorphism of M is conjugate to a pseudo Anosov map.*

Proof. By the previous lemma and Thurston's theorem [11] such an homeomorphism is isotopic to a Pseudo Anosov map.

References

1. Epstein, D., *Curves on 2-manifolds and isotopies*, Acta Mathematica **115** (1966), 83-107.
2. Fathi, A.; Laudenbach, F. and Poenaru, V., *Travaux de Thurston sur les surfaces*, (Seminaire Orsay), Asterisque No. **66-67** (1979).
3. Franks, J., *Anosov diffeomorphisms*, Proceedings of the Symposium in Pure Mathematics

14 (1970), 61-94.
4. Gerber, M. and Katok, A., *Smoth models of Thurston pseudo-Anosov maps*, Annales Scientifiques Ecole Normale Superieure **15** (1982), 173-204.
5. Handel, M., *Global shadowing of pseudo Anosov homeomorphisms*, Ergodic Theory and Dynamical Systems **5** (1985), 373-377.
6. Kuratowski, K., "Topology," Academic Press, New York-London, 1966.
7. Lewowicz, J., *Lyapunov functions and topological stability*, Journal of Differential Equations **38(2)** (1980), 192-209.
8. ________, *Persistence in expansive systems*, Ergodic Theory and Dynamical Systems **3** (1983), 567-578.
9. Lewowicz, J. and Lima de Sá, E., *Analytic models of pseudo Anosov maps*, Ergodic Theory and Dynamical Systems **6** (1986), 385-392.
10. O. Brien, T. and Reddy, W., *Each compact orientable surface of positive genus admits an expansive homeomorphism*, Pacific Journal Mathematics **35** (1970), 737-741.
11. Thurston, W., *On the geometry and dynamics of diffeomorphisms of surfaces*, Preprint.

Jorge Lewowicz
Instituto de Matemática y Estadística
Prof. Ing. Rafael Laguardia
Montevideo, Uruguay