Ergod. Th. & Dynam. Sys. (1990), **10**, 531-554
Printed in Great Britain

On Cherry flows

MARCO MARTENS† AND SEBASTIAN VAN STRIEN‡
Mathematics Department, Technical University of Delft, Julianalaan 132,
2628 *BL Delft, The Netherlands*

WELINGTON DE MELO
IMPA, Estrada Dona Castorina 110, 22460 *Rio de Janeiro RJ, Brazil*

PEDRO MENDES*
Departamento de Matemática, ICEx-UFMG, Cidade Universitária, Pampulha,
30000 *Belo Horizonte MG, Brazil*

(*Received* 1 *June* 1988 *and revised* 28 *July* 1989)

Abstract. The purpose of this research is to describe all smooth vector fields on the torus T^2 with a finite number of singularities, no periodic orbits and no saddle-connections. In this paper we are able to complete the description within the class of vector fields which are area contracting near all singularities. In particular we give a large class of analytic vector fields on the torus T^2 which have non-trivial recurrence and also sinks.

This result for vector fields follows from a result dealing with continuous monotone circle mappings which are possibly constant on a finite number of intervals.

1. *Statement of results for vector fields on T^2*

In this paper we are interested in recurrence of vector fields on surfaces. From the Poincaré-Bendixson theorem the dynamics of vector fields on the sphere is clear: the ω-limit of every point is equal to a point or to a closed orbit. On the torus T^2 the situation is more interesting. If a vector field on T^2 has periodic orbits then there are again no non-trivial recurrent orbits, because then the situation reduces to a planar one.

In his celebrated 1932 paper, Denjoy studied vector fields on the torus T^2 *without singularities* and *without periodic orbits*. He showed that if a vector field X on the torus is C^∞, then such a vector field either has a periodic orbit or every orbit of X is dense in T^2. In 1963, A. J. Schwartz extended Denjoy's result and showed that any minimal set of a C^∞ vector field on a compact connected, two dimensional manifold M is either a singleton (consisting of a singularity), or a single closed orbit, or all of M in which case $M = T^2$. (Here a set L is called *minimal* if it is a

† Supported by Netherlands Organisation for the Advancement of Pure Research (N.W.O.).
‡ Since most of the results in this paper were proved independently by M.M. & S.v.S and by W. de M. & P.M. it was decided to publish this paper jointly.
* Partially supported by the Brazilian CNPq, grants 300376/79/0 and 150013/87/MA/Fv.

closed, non-empty set which is invariant under the vector field such that L contains no smaller set with these properties.) From this it follows that the ω-limit of any point is equal to a periodic orbit, to $M = T^2$ or contains a singularity. But because no minimal set can strictly contain a singularity, the result of Schwartz gives a rather incomplete description of the dynamics of vector fields with singularities.

The next simplest vector fields on the torus have

(i) a finite number of singularities all of which are hyperbolic;
(ii) no periodic orbits;
(iii) no saddle-connections;
(iv) no sources.

We shall call a vector field which satisfies properties (i)-(iv) a Cherry vector field, because Cherry gave in 1938 an example of an analytic vector field on T^2 satisfying these properties with a sink and also a recurrent orbit. His example is quite specific, but shows that Cherry vector fields can have recurrent behaviour.

In this paper we will generalise Denjoy's result to all Cherry vector fields which have negative divergence at each saddle-point. This last condition is unnecessary if X is a Cherry vector field with at most one saddle-point. In this case, index arguments give that X can have only one other singularity. We may assume that the divergence in these singularities is of opposite signs, because otherwise either X or $-X$ would satisfy the conditions from before. So in this case we can assume that X has precisely one sink, no sources and one saddle-point. At this saddle-point the divergence of X is positive. This situation can be treated using the result of C. Gutierrez, A. Lins and W. de Melo described in § 6.

Theorem A implies that there exists a large class of vector fields which display the same phenomenon as Cherry's example.

Let $\mathcal{B}$ be the class of C^∞ Cherry vector fields X on T^2 such that at each singularity p the divergence of X is smaller or equal to 0. Furthermore, let $\mathcal{B}'$ be the class of C^∞ Cherry vector fields X on T^2 with precisely one hyperbolic sink and one saddle-point; moreover assume that X has positive divergence at this saddle-point. (As we noted above, for each vector field X with only hyperbolic singularities and one saddle-point either $X \in \mathcal{B} \cup \mathcal{B}'$ or $-X \in \mathcal{B} \cup \mathcal{B}'$.)

Denote the flow of X through x by $t \to X_t(x)$. The α and ω limit set of x are defined as $\alpha(x) = \{y; \exists t_n \to \infty$ with $X_{-t_n}(x) \to y\}$, respectively $\omega(x) = \{y; \exists t_n \to \infty$ with $X_{t_n}(x) \to y\}$. We say that x is *recurrent* if $x \in \alpha(x) \cup \omega(x)$.

THEOREM A. *Let $X \in \mathcal{B} \cup \mathcal{B}'$. Then there exists an infinite set $L \subset T^2$ such that every point $x \in L$ is recurrent, $\omega(x) = \alpha(x) = L$. Moreover, for every $x \in T^2$ which is not contained in the unstable manifold of a saddle point one has $\alpha(x) = L$ and for every point $x \in T^2$ which is not contained in the stable manifold of a sink or of a saddle, $x \in L$.*

Remark 1. The condition that $X \in \mathcal{B} \cup \mathcal{B}'$ implies that X cannot have any sources.

Remark 2. It is not very difficult to construct many vector fields which satisfy the assumptions of Theorem A.

Remark 3. Of course Denjoy's result also holds for C^2 vector fields. Our result also holds for C^2 vector fields, but in this case we have to require that there exist C^2

linearising coordinates near every saddle-point of X, or alternatively that X is C^6 near saddle-points (see the Appendix).

The next result says that vector fields $X \in \mathcal{B}$ can be described by circle maps.

THEOREM B. *Let $X \in \mathcal{B} \cup \mathcal{B}'$. Then there exists a neighbourhood U of X such that for every $Y \in U$ one can associate a continuous order-preserving return-map $f_Y : S^1 \to S^1$ with the following properties:*
(a) *f_Y depends continuously on $Y \in U$;*
(b) *if Y, $Z \in U$ have no periodic orbits and no saddle-connections then they are equivalent if f_Y and f_Z are topologically conjugate circle maps.*

In fact, these circle maps may be constant on some intervals. Also they will satisfy some smoothness conditions and some non-flatness conditions at the boundary of the intervals where they are constant. In the next section we will study a special class $\mathcal{A}$ of circle mappings and show that if these maps have no periodic points then they are determined by a finite number of parameters, see Theorem D. In § 6 we will show that the maps f_X are contained in this class $\mathcal{A}$, provided some transition map is sufficiently smooth and non-flat. In the Appendix we will show that these conditions hold for all smooth vector fields. (These conditions would follow immediately if we assumed that X was C^2 linearizable at singularities. However, it is rather pretty to see how natural the conditions that one obtains without assuming linearizability correspond to the conditions needed in the class $\mathcal{A}$.)

2. *Statements of results for circle maps*

Let $S^1 = \mathbb{R} (\text{mod } 1)$ and $f: S^1 \to S^1$ be a continuous, order-preserving, degree one map which is possibly constant on a finite number of intervals. Since f is order preserving it has a well-defined rotation number $\rho(f) \in [0, 1)$. $\rho(f)$ is irrational if and only if f has no periodic points. It is well known that if $\rho(f)$ is irrational then f is semi-conjugate to the rotation $R_{\rho(f)}: x \to x + \rho(f)$ mod 1. The semi-conjugacy is continuous, monotone, has degree one and maps orbits of f to orbits of $R_{\rho(f)}$.

Assume that $f: S^1 \to S^1$ is continuous, order-preserving, degree one and everywhere C^1 except possibly in a finite number of points. Let K_f be the closure of the set of points x with either
- f is not C^1 in x, or
- f is constant on some neighbourhood of x, or
- $Df(x) = 0$.

Let C_f be the closure of the set of points where f is locally constant.

We say that such a map f is in $\mathcal{A}$ if f satisfies the conditions from above, K_f has finitely many components and also the following smoothness and non-flatness conditions. For each point $x_0 \in S^1 \setminus K_f$ there exists a neighbourhood U of x_0 such that f is a local C^1 diffeomorphism on U and Df has bounded variation on U. f is constant on each component of K_f. (In particular f is everywhere C^1 except possibly in points $x \in \partial K_f$.) Moreover we require that for every point $x_0 \in \partial K_f$, f satisfies the following non-flatness condition. There exist neighbourhoods U, V of

https://doi.org/10.1017/S0143385700005733 Published online by Cambridge University Press

x_0 respectively $f(x_0)$, a constant $\alpha \geq 1$, and two C^1 diffeomorphisms $\phi_r : ((-1, 1), 0) \to (U, x)$ and $\phi_l : (f(U), f(x)) \to ((-1, 1), 0)$ such that $D\phi_l$ and $D\phi_r$ have bounded variation and $\phi_l \circ f \circ \phi_r(x)$ is equal to 0, or $\pm|x|^\alpha$.

Similarly we say that $f \in \mathcal{A}'$ if f has degree one, is everywhere strictly increasing except on one interval K_f where it is constant and $f|(S^1 - K_f)$ is C^1. Moreover at the boundary points of $K_f = [a, b]$ the (one-sided) derivative of f is infinite and there exists $\varepsilon > 0$ such that $\log Df|[a-\varepsilon, a)$ and $\log Df|(b, b+\varepsilon]$ are monotone.

Remark. If f is C^2 on $S^1 \backslash K_f$, and the limits $\lim_{x \to x_0, x \notin K_f} D^2 f(x)$ are non-zero for each point $x_0 \in \partial K_f$, then $f \in \mathcal{A}$. However, we will not restrict ourselves to such maps, because, in applying our results to vector fields on the torus, we get that singularities of the vector field of saddle-type give rise to critical points of a corresponding circle map of the form $|x|^\delta$, where $\delta > 0$ and is the absolute value of the ratio of the eigenvalues at the saddle-point and need not be an integer.

We can also consider the slightly more general class of maps such that there exists neighbourhoods $V \subset \text{Clos}(V) \subset U$ of K_f such that the maps $f|(S^1 \backslash V)$, $\phi_l | U$ and $\phi_r | U$ are absolutely continuous and the maps $\log Df|(S^1 \backslash V)$, $\log D\phi_l | U$ and $\log D\phi_r | U$ (which therefore exist almost everywhere) are almost everywhere equal to maps with bounded variation. The proofs of our results go through without much change; we indicate the changes needed in the remark following Proposition 3.4.

We say that I is a *wandering interval* of $f : S^1 \to S^1$ if

- $f^n(I) \cap f^m(I) = \emptyset$, $\forall n, m \geq 0$ with $n \neq m$;
- there exists no $n \geq 0$ such that $f^n(I) \subset C_f$.

THEOREM C. *Let f be in $\mathcal{A} \cup \mathcal{A}'$ with irrational rotation number. Then f has no wandering intervals.*

COROLLARY. *For every $x \in S^1 \backslash \bigcup_{n \geq 0} f^{-n}(C_f)$ both the forward and the backward orbit of x is dense in $S^1 \backslash \bigcup_{n \geq 0} f^{-n}(C_f)$. In particular there exists a non-finite recurrent orbit of f.*

Proof of Corollary. Since f has no wandering intervals, the semi-conjugacy h_f between f and $R_{\rho(f)}$ is constant only on preimages of C_f. Since $S^1 \backslash \bigcup_{n \geq 0} f^{-n}(C_f)$ cannot be empty the Corollary follows. □

Remark 1. J. C. Yoccoz has proved Theorem C for maps f which are strictly monotone (isolated critical points are still allowed). However, if a vector field on T^2 has sinks then the corresponding return-map is constant on some interval. So the result of Yoccoz cannot be applied to vector fields. Unfortunately the proof of his result cannot be directly generalized to maps which are allowed to be constant on some intervals, see the remarks below: Definition 3.1, Lemma 3.7 and in § 5.

Remark 2. Consider order-preserving maps $f : S^1 \to S^1$ of degree one which are constant on at least one interval and which are discontinuous in at least one point. Let K_f be the closure of the set of points where f is locally constant or is discontinuous. Assume that f satisfies the same smoothness conditions as the maps from $\mathcal{A}$ (except that f is not continuous). Then Gutierrez has shown that either $\rho(f)$ is rational or some iterate of one of the components of K_f is mapped into some other component of K_f.

So the main thing which is missing in the general description of these maps is the case where $\alpha > 1$ in some points in ∂K_f and $\alpha < 1$ in other points of ∂K_f.

From this result one can easily show that the topological conjugacy class of f is determined by a finite number of invariants. Indeed, if $\rho(f)$ is irrational, then there exists an order-preserving map $h_f : S^1 \to S^1$ such that $h_f \circ f = R_\rho \circ h_f$, where $R_\rho(x) = x + \rho \pmod 1$ and $\rho = \rho(f)$, see [**He**]. For simplicity write $R = R_\rho$. Clearly h_f is constant on the components of C_f.

Take $f \in \mathscr{A} \cup \mathscr{A}'$. If $\rho(f)$ is irrational and h_f is constant on some interval I then $R^n(h_f(I))$ is disjoint for all $n \geq 0$. Hence the intervals $f^n(I)$ are mutually disjoint for all n and from Theorem C there exists $n \geq 0$ with $f^n(I) \subset C_f$. It follows that $h_f^{-1}(h_f(x))$ is a non-trivial interval if and only if $x \in \bigcup_{n \geq 0} f^{-n}(C_f)$. Let d be the metric on $S^1 = \mathbb{R} \pmod 1$ induced by $\mathbb{R}$.

THEOREM D. *Let $f, g \in \mathscr{A} \cup \mathscr{A}'$ with irrational rotation number such that the number of components of C_f is the same as that of C_g. Call this number k. Assume that*

$$\rho(f) = \rho(g) \tag{2.1}$$

and that we can order the components C_f^i and C_g^i of C_f respectively C_g cyclically on S^1 such that

$$d(c_f^i, c_f^{(i+1)}) = d(c_g^i, c_g^{(i+1)}), \quad i = 1, 2, \ldots, k-1. \tag{2.2}$$

Here $c_f^i = h_f(C_f^i)$ and $c_g^i = h_g(C_g^i)$. Then f and g are conjugate, i.e. there exists a homeomorphism $h : S^1 \to S^1$ such that $h \circ f = g \circ h$.

Proof of Theorem D. We begin by taking an arbitrary order-preserving homeomorphism $k : C_f \to C_g$ such that $k(C_f^i) = C_g^i$. Then, let h_f and h_g be the order-preserving maps such that $h_f \circ f = R_\rho \circ h_f$ and $h_g \circ g = R_\rho \circ h_g$. Choosing θ so that R_θ maps $h_f(C_f^1)$ onto $h_g(C_g^1)$, one gets from (2.2) that $R_\theta(h_f(C_f)) = h_g(C_g)$. Moreover taking $\tilde{h}_f = R_\theta \circ h_f$, the following diagram commutes

$$\begin{array}{ccc} S^1 & \xrightarrow{\ f\ } & S^1 \\ {\scriptstyle \tilde{h}_f}\downarrow & & \downarrow{\scriptstyle \tilde{h}_f} \\ S^1 & \xrightarrow{\ R_\rho\ } & S^1 \\ {\scriptstyle h_g}\uparrow & & \uparrow{\scriptstyle h_g} \\ S^1 & \xrightarrow{\ g\ } & S^1 \end{array}$$

and that $\tilde{h}_f(C_f^i) = C_g^i$. So

$$h(x) = \begin{cases} h_g^{-1} \circ \tilde{h}_f(x), & \text{for } x \notin \bigcup_{n \geq 0} f^{-n}(C_f) \\ k(x), & \text{for } x \in C_f, \end{cases}$$

https://doi.org/10.1017/S0143385700005733 Published online by Cambridge University Press

defines a strictly order-preserving homeomorphism from $S^1\backslash\bigcup_{n\geq 1} f^{-n}(C_f)$ to $S^1\backslash\bigcup_{n\geq 1} g^{-n}(C_g)$ such that $h\circ f=g\circ h$. Finally, for $x\in\bigcup_{n\geq 1} f^{-n}(C_f)$, let $n\geq 1$ be the minimal number such that $f^n(x)\in C_f$ and define $h(x)=g^{-n}\circ h\circ f^n$. In this way we have defined a homeomorphism h such that $h\circ f=g\circ h$. □

In the Proof of Theorem C, we will need a simple Proposition which produces collections of disjoint intervals. Since it does not fit logically in the next two sections we state and prove it here.

PROPOSITION 2.1. *Let $\{W_i|i=1,\ldots,m\}$ be a collection of intervals in S^1 with the following two properties.*
(i) *every point in S^1 is in at most three elements of the collection;*
(ii) *$W_i\subset W_k$ implies $i=k$.*
Then there exists a partition of the collection

$$\{W_i;\, i=0,1,\ldots,m\}=A_1\cup A_2\cup\cdots\cup A_5,$$

such that A_k consists of mutually disjoint intervals for each $k=1,2,\ldots,5$.

Proof. We prove this Proposition inductively for each $m\geq 1$. The Proposition is trivially true for $m=1$. Suppose that it is proved inductively for $m-1$. Then we can write

$$\{W_i;\, i=1,2,\ldots,m-1\}=\bigcup_{j=1,2,\ldots,5} A'_j$$

where A'_j consists of mutually disjoint intervals. From property (ii) it follows that $\partial W_m\cap W_i\neq\emptyset$ when $W_m\cap W_i\neq\emptyset$. So assumption (i) implies that there exists at most four is with $1\leq i<m$ such that W_m has a non-empty intersection with W_i. Hence there exists $j'\in\{1,2,\ldots,5\}$ such that W_m does not intersect any interval from $A'_{j'}$. So by letting $A_j=A'_j$ for $j\in\{1,2,\ldots,5\}$, $j\neq j'$ and $A_{j'}=A'_{j'}\cup\{W_m\}$ we get the required partition of $\{W_i;\, i=1,2,\ldots,m\}$. This proves the Proposition by induction. □

The organization of the paper is as follows. In the next three sections we will prove Theorem C. In § 6 we will prove Theorems A and B, except a technical Lemma which is proved in an Appendix.

It will be convenient to introduce the following notation. For an interval I let $|I|$ denote the length of I. Moreover let d be the metric on $S^1=\mathbb{R}\pmod 1$ induced by $\mathbb{R}$.

3. *Some analytic tools*

In this section we will develop some analytic tools which give estimates on the shape of non-linearity of f^n on intervals T such that $T, f(T),\ldots,f^n(T)$ are disjoint.

Definition 3.1. Let $g: S^1\to S^1$ be a continuous order-preserving map and I, T intervals in S^1 with $\text{Clos}(I)\subset\text{int}(T)$. Then

$$D(T,I)=\frac{|T|\cdot|I|}{|L|\cdot|R|},\ \text{where } L \text{ and } R \text{ are the components of } T\backslash I;$$

$$B(g,T,I)=\frac{D(g(T),g(I))}{D(T,I)}.$$

Remark. J. C. Yoccoz uses the operator $B_0(g, T) = \lim_{I \to T} B(g, T, I)$.

Definition 3.2. Let $g: T \to S^1$ be a C^3-map on the interval $T \subset S^1$. Then the operator

$$Sg(x) = \frac{g'''(x)}{g'(x)} - \frac{3}{2}\left(\frac{g''(x)}{g'(x)}\right)^2$$

is called the Schwarzian derivative.

It is easy to see that the operator B is multiplicative:

$$B(g \circ h, T, I) = B(g, h(T), h(I)) \times B(h, T, I). \tag{3.1}$$

The proof of the following proposition can be found in [**MS**].

PROPOSITION 3.3. *Let $g: T \to T$ be a C^3 diffeomorphism on the open interval T such that for every $x \in T$, $Sg(x) < 0$. Let* $\mathrm{Clos}(I) \subset \mathrm{int}(T)$. *Then*

$$B(g, T, I) > 1. \tag{3.2}$$

Let $\mathcal{V} \subset \mathrm{Clos}(\mathcal{V}) \subset \mathrm{int}\,(\mathcal{U})$ be interval neighbourhoods of K_f such that the number of components of $\mathcal{V}$ and $\mathcal{U}$ are both the same as $\# K_f$.

PROPOSITION 3.4. *Let $f \in \mathcal{A}$. Then there exists $V < \infty$ with the following property. Let $\Xi = \{T_1, T_2, \ldots, T_n\}$ be a collection consisting of n intervals in S^1 with the following properties. Every point of S^1 is contained in at most three intervals T_i, $T_i \subset T_j$ implies $i = j$ and none of the intervals T_i contains a component of $\mathcal{U} \backslash \mathcal{V}$ or contains points of K_f. Then*

$$\sum_{i=1}^{n} \log B(f, T_i, I_i) \geq -5 \cdot V. \tag{3.3}$$

Proof. Let $V_f = \mathrm{Var}\,(\log Df \,|\, (S^1 \backslash \mathcal{V}))$. Since on $S^1 \backslash \mathcal{V}$ the map f is a C^1 diffeomorphism and Df has bounded variation, $V_f < \infty$. Moreover on each component $\mathcal{U}_i$ of $\mathcal{U}$, f has the form $f(x) = \phi_{l,i} \circ \phi_\alpha \circ \phi_{r,i}$, where $\phi_\alpha(x) = \pm |x|^\alpha$ as above. Here $\phi_{l,i}$ and $\phi_{r,i}$ are C^1 diffeomorphisms such that $\log D\phi_{l,i}$ and $\log D\phi_{r,i}$ have bounded variations on $\mathcal{U}_i$. Hence $V_l = \sum_i \mathrm{Var}\,(\log\,(D\phi_{l,i}))$ and $V_r = \sum_i \mathrm{Var}\,(\log\,(D\phi_{r,i}))$ are both finite. Let $V = V_f + V_l + V_r$.

Let $I_1 = \{i;\, T_i \cap \mathcal{V} = \emptyset\}$ and $I_2 = \{i;\, T_i \subset \mathcal{U}\}$. Since the intervals T_i never contain a component of $\mathcal{U} \backslash \mathcal{V}$, $I_1 \cup I_2 = \{1, 2, \ldots, n\}$.

First assume that $i \in I_1$. Let L_i and R_i be the components of $T_i \backslash I_i$. For $u_i, v_i \in T_i$ let (u_i, v_i) be the open interval connecting u_i and v_i. Because f is C^1 on $S^1 \backslash \mathcal{V}$ the mean-value theorem is valid on T_i. Using this in the definition of $B(f, T_i, I_i)$ we find that there exist $l_i \in L_i$, $r_i \in R_i$, $m_i \in I_i$ and $\tau_i \in T_i$ such that

$$\log B(f, T_i, I_i) = \log\left(\frac{Df(m_i)Df(\tau_i)}{Df(l_i)Df(r_i)}\right), \tag{3.4}$$

and

$$m_i \in (l_i, r_i). \tag{3.5}$$

From (3.4) one has

$$|\log B(f, T_i, I_i)| \le |\log Df(m_i) - \log Df(l_i)| + |\log Df(\tau_i) - \log Df(r_i)| \quad (3.6)$$

and also

$$|\log B(f, T_i, I_i)| \le |\log Df(m_i) - \log Df(r_i)| + |\log Df(\tau_i) - \log Df(l_i)|. \quad (3.7)$$

Rename the points l_i, m_i, r_i, τ_i in increasing order a_i, b_i, c_i, d_i. From (3.5) one gets that either $(l_i, m_i) \cap (\tau_i, r_i) = \emptyset$ or $(\tau_i, l_i) \cap (m_i, r_i) = \emptyset$, and so we can use either (3.6) or (3.7) and get

$$\begin{aligned}|\log B(f, T_i, I_i)| &\le |\log Df(b_i) - \log Df(a_i)| + |\log Df(d_i) - \log Df(c_i)| \\ &\le \mathrm{Var}\,(\log Df \,|\, T_i)\end{aligned}$$

and therefore

$$\log B(f, T_i, I_i) \ge -\mathrm{Var}\,(\log Df \,|\, T_i). \quad (3.8)$$

Now consider $i \in I_2$. Then T_i is contained in some component $\mathcal{U}_i$ of $\mathcal{U}$ (and does not intersect K_f) and so f has the form $f(x) = \phi_{l,i} \circ \phi_\alpha \circ \phi_{r,i}$. Hence

$$B(f, T_i, I_i) = B(\phi_{l,i}, T_i'', I_i'') \times B(\phi_\alpha, T_i', I_i') \times B(\phi_{r,i}, T_i, I_i).$$

Here $T_i'' = \phi_\alpha \circ \phi_{r,i}(T_i)$, $T_i' = \phi_{r,i}(T_i)$, $I_i'' = \phi_\alpha \circ \phi_{r,i}(I_i)$ and $I_i' = \phi_{r,i}(I_i)$. Since the Schwarzian derivative of ϕ_α is less or equal to 0 (because $\alpha \ge 1$) one gets $B(\phi_\alpha, T_i', I_i') \ge 1$. Hence, as above,

$$\begin{aligned}\log B(f, T_i, I_i) &\ge \log B(\phi_{l,i}, T_i'', I_i'') + 0 + \log B(\phi_{r,i}, T_i, I_i) \\ &\ge -\{\mathrm{Var}\,(\log D\phi_{l,i} | T_i'') + \mathrm{Var}\,(\log D\phi_{r,i} | T_i)\}. \quad (3.9)\end{aligned}$$

Since every point of S^1 is contained in at most three intervals, using Proposition 2.1, one can write $\{T_i; i = 1, 2, \dots, n\} = A_1 \cup A_2 \cdots \cup A_5$ where A_j consists of a collection of mutually disjoint intervals. Therefore also the collection of intervals T_i'' corresponding to $T_i \in A_j$ consists of disjoint intervals. Hence from (3.8) and (3.9) one gets

$$\sum_{i=1}^{n} \log B(f, T_i, I_i) \ge -5 \cdot (V_f + V_l + V_r).$$

The Proposition follows. □

Remark. In eq. (3.4) in this Proposition we have used that Df exists and has bounded variation on $S^1 \setminus \mathcal{V}$. Of course it would have been sufficient to assume that f is absolutely continuous, $\log Df | (S^1 \setminus \mathcal{V})$ is almost everywhere equal to a map with bounded variation. In this case we could have still found $m_i \in I_i$, $l_i \in L_i$, $r_i \in R_i$ and $\tau_i \in T_i$ such that Df exists in these points and such that

$$\frac{|f(I_i)|}{|I_i|} \le Df(m_i), \quad \frac{|f(T_i)|}{|T_i|} \le Df(\tau_i), \quad \frac{|f(L_i)|}{|L_i|} \ge Df(l_i) \quad \text{and} \quad \frac{|f(R_i)|}{|R_i|} \ge Df(r_i).$$

Hence (3.6), (3.7) and (3.8) would still hold. Similarly it suffices to assume that ϕ_l, ϕ_r are absolutely continuous and $\log D\phi_l$ and $\log D\phi_r$ are almost everywhere equal to maps with bounded variation.

https://doi.org/10.1017/S0143385700005733 Published online by Cambridge University Press

LEMMA 3.5. *Let $f \in \mathcal{A}$. Then there exists $A_0 > 0$ such that if $I \subset W$ are intervals in S^1 such that*

(a) *$W \backslash I$ consists of one component H;*

(b) $|H| \leq |I|$;

(c) $I \cap K_f = \emptyset$,

then

$$\frac{|f(I)|}{|f(H)|} \geq A_0 \cdot \frac{|I|}{|H|}. \tag{3.10}$$

Proof. Let us prove this Proposition by contradiction. So suppose that there exists a sequence of intervals I_n, W_n, $H_n = W_n \backslash I_n$ satisfying (a), (b) and (c) above such that

$$\lim_{n \to \infty} \frac{|f(I_n)|/|I_n|}{|f(H_n)|/|H_n|} \to 0. \tag{3.11}$$

Since $|H_n| \leq |I_n|$ this implies that $|W_n| \to 0$ and $W_n \to K_f$. By taking a subsequence we may assume that $W_n \to p \in K_f$.

Since $f \in \mathcal{A}$ we may assume that p is a critical point and near p, f is of the form $f = \phi_l \circ \phi \circ \phi_r$ where ϕ_α is equal to x^α or 0 on $x \geq 0$ and to 0 or $-|x|^\alpha$ on $x \leq 0$ as before. Hence

$$\frac{|f(I_n)|/|I_n|}{|f(H_n)|/|H_n|} \geq \frac{\min D\phi_l}{\sup D\phi_l} \times \frac{\min D\phi_r}{\sup D\phi_r} \times \frac{|\phi_\alpha(I'_n)|/|I'_n|}{|\phi_\alpha(H'_n)|/|H'_n|}. \tag{3.12}$$

Here $H'_n = \phi_r(H_n)$ and $I'_n = \phi_r(I_n)$. Since ϕ_r is a C^1 diffeomorphism, writing $K_n = |H'_n|/|I'_n|$ we get $\limsup K_n < \infty$ and therefore there exists $K \in (0, \infty)$ such that $K_n \leq K$, $\forall n \geq 0$. Let us distinguish between the following three possibilities.

Case 1. I_n is between p and H_n.

Writing $I'_n = [a_n, b_n]$, $H'_n = [b_n, c_n] \subset (0, 1)$ one has

$$\frac{|\phi_\alpha(I'_n)|/|I'_n|}{|\phi_\alpha(H'_n)|/|H'_n|} = \frac{\left[\dfrac{b_n^\alpha - a_n^\alpha}{b_n - a_n}\right]}{\left[\dfrac{c_n^\alpha - b_n^\alpha}{c_n - b_n}\right]} \geq \frac{\left[\dfrac{b_n^\alpha}{b_n}\right]}{\left[\dfrac{c_n^\alpha - b_n^\alpha}{c_n - b_n}\right]} \geq \frac{\left[\dfrac{b_n^\alpha}{b_n}\right]}{\left[\dfrac{((1+K)b_n)^\alpha - (b_n)^\alpha}{(1+K)b_n - b_n}\right]}$$

$$= \frac{K}{(1+K)^\alpha - 1} > 0. \tag{3.13}$$

Here, the second inequality follows because $\alpha \geq 1$ implies that the function $x \to Dx^\alpha = \alpha \cdot x^{\alpha - 1}$ is increasing. The third inequality follows from this and because $|H'_n| \leq K'_n \cdot |I_n|$ gives

$$c_n \leq b_n + K_n \cdot (b_n - a_n) \leq b_n + K_n \cdot b_n \leq (1+K) \cdot b_n.$$

Using (3.13) in (3.12) gives a contradiction in (3.11).

Case 2. H_n is between p and I_n, but $H_n \cap K_f = \emptyset$.

Since the derivative of ϕ_α is increasing we have in this case

$$\frac{|\phi_\alpha(I'_n)/|I'_n|}{|\phi_\alpha(H'_n)|/|H'_n|} \geq 1. \tag{3.14}$$

Using (3.14) in (3.12) we again get a contradiction with (3.11).

Case 3. H_n is between p and I_n, but $H_n \cap K_f \neq \varnothing$.

If H_n intersects an interval component of K_f, then we can choose the maximal interval $\tilde{W}_n$ with $I_n \subset \tilde{W}_n \subset W_n$ such that $\tilde{W}_n \cap K_f = \varnothing$. Let $\tilde{H}_n = \tilde{W}_n \backslash I_n$ and $\tilde{H}'_n = \phi_r(H'_n)$. By construction $|\tilde{H}'_n| \leq |H'_n|$ and $\phi_\alpha(\tilde{H}'_n) = \phi_\alpha(H'_n)$, and therefore

$$\frac{|\phi_\alpha(I'_n)|/|I'_n|}{|\phi_\alpha(H'_n)|/|H'_n|} \geq \frac{|\phi_\alpha(I'_n)|/|I'_n|}{|\phi_\alpha(\tilde{H}'_n)|/|\tilde{H}'_n|}.$$

As in case 2 this last term is bounded away from zero and we get a contradiction as before.

So we may assume that $H_n \cap K_f$ is an isolated point. So writing $H'_n = [-a_n, b_n]$, $I'_n = [b_n, c_n]$, with $a_n, b_n, c_n > 0$ one has

$$\frac{|\phi_\alpha(I'_n)|/|I'_n|}{|\phi_\alpha(H'_n)/|H'_n|} = \frac{\left[\dfrac{c_n^\alpha - b_n^\alpha}{c_n - b_n}\right]}{\left[\dfrac{b_n^\alpha + a_n^\alpha}{b_n + a_n}\right]} \geq \frac{\left[\dfrac{c_n^\alpha - b_n^\alpha}{c_n - b_n}\right]}{\left[\dfrac{b_n^\alpha}{b_n} + \dfrac{a_n^\alpha}{a_n}\right]} \geq \frac{\left[\dfrac{c_n^\alpha}{c_n}\right]}{\left[\dfrac{c_n^\alpha}{c_n} + \dfrac{a_n^\alpha}{a_n}\right]} \geq \frac{1}{(1+K)^\alpha}. \tag{3.15}$$

Here the first inequality follows from

$$\frac{x+u}{y+v} \leq \frac{x}{y} + \frac{u}{v} \quad \text{for all } x, y, u, v > 0.$$

The last three inequalities follow from the concavity of x^α and because $b_n, a_n \leq |H'_n| \leq K \cdot |I'_n| \leq K \cdot c_n$. So using (3.15) in (3.12) we get again a contradiction with (3.11).

Since cases 1, 2 and 3 cover all possibilities we have proved the Proposition by contradiction. □

PROPOSITION 3.6. *Let $f \in \mathscr{A}$ and I, T be intervals in S^1 with* Clos$(I) \subset$ int(T) *and L and R be the components of $T \backslash I$. Let $\gamma \in (0, 1)$. If*

$$\frac{|f(I)|}{|f(R)|} \geq \gamma \cdot \frac{|I|}{|R|} \quad \text{and} \quad \frac{|f(I)|}{|f(L)|} \geq \gamma \cdot \frac{|I|}{|L|},$$

then

$$B(f, T, I) \geq \gamma^2.$$

Proof.

$$B(f, T, I) = \frac{\dfrac{|f(I)|}{|f(R)|} + \dfrac{|f(I)|}{|f(R)|}\dfrac{|f(I)|}{|f(L)|} + \dfrac{|f(I)|}{|f(L)|}}{\dfrac{|I|}{|R|} + \dfrac{|I|}{|R|}\dfrac{|I|}{|L|} + \dfrac{|I|}{|L|}} \geq \frac{\gamma + \gamma^2 \cdot d}{1+d} = \gamma \cdot \frac{1+\gamma \cdot d}{1+d} \geq \gamma^2.$$

Here

$$d = \frac{\dfrac{|I||I|}{|R||L|}}{\dfrac{|I|}{|R|} + \dfrac{|I|}{|L|}}$$

and the last inequality follows since $\gamma \in (0,1)$ and therefore $(0,\infty) \ni x \to (1+\gamma x)/(1+x)$ is always greater than γ. The result follows. □

PROPOSITION 3.7. *Let $f \in \mathcal{A}$. Then there exists $A_1 > 0$ such that if I, T are intervals in S^1 with* $\text{Clos}(I) \subset \text{int}(T)$ *and L and R the components of $T \setminus I$ such that*
(a) $|L| \le |I|$ *or* $|R| \le |I|$;
(b) $I \cap K_f = \varnothing$;
then

$$B(f, T, I) \ge A_1. \tag{3.16}$$

Proof. Let A_0 be the number from Proposition 3.5. We may assume that $A_0 \in (0,1)$. We will prove the Proposition for $A_1 = \frac{1}{3} \cdot (A_0)^2$. By possibly renaming L and R, we may consider the case that $|R| \le |I|$. Then from Proposition 3.5 we get

$$\frac{|f(I)|/|I|}{|f(R)|/|R|} \ge A_0, \tag{3.17}$$

and hence

$$B(f, T, I) \ge \frac{|f(T)|}{|f(L)|} \cdot \frac{|L|}{|T|} \cdot A_0 \ge A_0 \cdot \frac{|L|}{|T|}.$$

If $|L| \ge |I|$ then it follows from this and $|R| \le |I|$ that

$$B(f, T, I) \ge A_0 \cdot \frac{|L|}{|L|+|I|+|R|} \ge A_0 \cdot \frac{|I|}{2|I|+|R|} \ge A_0 \cdot \tfrac{1}{3} \ge A_1$$

and the Proposition is proved. So assume that $|L| \le |I|$. Then applying Proposition 3.5 again we get

$$\frac{|f(I)|/|I|}{|f(L)|/|L|} \ge A_0, \tag{3.18}$$

and it follows from (3.17), (3.18) and Proposition 3.6 that $B(f, T, I) \ge A_0 \cdot A_0 = A_1$. □

Remark. In general the interval T we need to consider may intersect or even overlap a component of K_f. It is easy to check that there exists no universal lower bound for $B_0(f, T)$ (as in (3.16)) in this case. Here B_0 is the operator of Yoccoz mentioned in the remark below Definition 3.1. Therefore we have to replace the operator B_0 by B.

Let $\mathcal{V} \subset \mathcal{U}$ be the neighbourhoods of K_f from above.

THEOREM 3.8. *Let $f \in \mathcal{A}$. Then there exists $\varepsilon > 0$ and B_0 such that for any intervals* $I \subset \text{Clos}(I) \subset \text{int}(T)$ *in S^1, and any $n \ge 0$, satisfying the following conditions*
(a) $|L| < \varepsilon \cdot |I|$ *and* $|R| < \varepsilon \cdot |I|$; *(Here L and R are the component of $T \setminus I$.)*
(b) *every point in S^1 is contained in at most three members of the sequence T, $f(T), \dots, f^{n-1}(T)$;*
(c) $f^i(I) \cap K_f = \varnothing$, $i \in \{0, 1, \dots, n-1\}$ *then*

$$B(f^n, T, I) \ge B_0. \tag{3.19}$$

Proof. Let A_1 be the number from Proposition 3.7 and assume that $A_1 < 1$ and $e^{-5V} < 1$. Let V be the number from Proposition 3.3. We may assume that $e^{-5V} < 1$.

Let $B_0=\{A_1\, e^{-5V}\}^{12\#K_f+1}$ and $\varepsilon=\sqrt{B_0/3}$. Let $0\le t_1<t_2<\cdots<t_s\le n-1$ be integers such that either $f^i(T)$ contains a component of $\mathcal{U}\backslash\mathcal{V}$, or $f^i(T)\cap K_f\neq\emptyset$. From the disjointness property of the orbit of T we get $s\le 12\#K_f$. From the definition of t_1 one has that $f^i(T)$ does not contain a component of $\mathcal{U}\backslash\mathcal{V}$ for $i<t_1$ and because f does not have periodic points we have that $f^i(T)\subset f^j(T)$ implies $i=j$. Using this and the disjointness property of $f^i(T)$, Proposition 3.4 gives

$$B(f^{t_1},T,I)\ge e^{-5V}.$$

So

$$\frac{|f^{t_1}(T)||f^{t_1}(I)|}{|f^{t_1}(L)||f^{t_1}(R)|}\ge e^{-5V}\cdot\frac{|T||I|}{|L||R|}\ge B_0\cdot\frac{|T||I|}{|L||R|}.$$

This implies

$$\frac{|f^{t_1}(I)|}{|f^{t_1}(R)|}+\frac{|f^{t_1}(I)|}{|f^{t_1}(R)|}\frac{|f^{t_1}(I)|}{|f^{t_1}(L)|}+\frac{|f^{t_1}(I)}{|f^{t_1}(L)|}\ge B_0\cdot\frac{1}{\varepsilon^2}=3.$$

So $|f^{t_1}(R)|\le|f^{t_1}(I)|$ or $|f^{t_1}(L)|\le|f^{t_1}(I)|$. Hence we can apply Proposition 3.7 and get $B(f,f^{t_1}(T),f^{t_1}(I))\ge A_1$ and hence

$$B(f^{t_1+1},T,I)\ge B(f,f^{t_1}(T),f^{t_1}(I))\times B(f^{t_1},T,I)\ge e^{-5V}\cdot A_1. \tag{3.20}$$

In the same way we get

$$B(f^{t_2-(t_1+1)},f^{t_1+1}(T),f^{t_1+1}(I))\ge e^{-5V},$$

and hence, using (3.20),

$$B(f^{t_2},T,I)\ge[e^{-5V}]^2\cdot A_1\ge B_0.$$

From this we get as before $|f^{t_2}(R)|\le|f^{t_2}(I)|$ or $|f^{t_2}(L)|\le|f^{t_2}(I)|$. Hence we can apply Proposition 3.7 again and get

$$B(f^{t_2+1},T,I)\ge[e^{-5V}\cdot A_1]^2\ge B_0.$$

This procedure has to be repeated at most $12\cdot\#K_f$ times, i.e., $s\le 12\cdot\#K_f$. Since $[e^{-5V}\cdot A_1]^{12\cdot\#K_f}\ge B_0$ we get by induction that for each $k\le s$, $|f^{t_k}(R)|\le|f^{t_k}(I)|$ or $|f^{t_k}(L)|\le|f^{t_k}(I)|$ and hence

$$B(f^{t_s+1},T,I)\ge[e^{-5V}\cdot A_1]^{12\cdot\#K_f}.$$

Using Proposition 3.4 we can handle the last piece from t_s+1 to n and get

$$B(f^{t_s+1},T,I)\ge[e^{-5V}\cdot A_1]^{12\cdot\#K_f}\cdot e^{-5V}.$$ □

4. *The topological situation*

Throughout this section we will assume that f has a wandering interval I. We may assume that I is a maximal wandering interval, i.e. not contained in a bigger wandering interval. Since the forward orbit of I is disjoint and K_f has at most a finite number of components, by possibly replacing I by some iterate of I we can assume that $f^i(I)\cap K_f=\emptyset$. We want to show that there exist iterates q_n and neighbourhoods T_n of I such that f^{q_n} is contracting on I but strongly expanding on $T_n\backslash I$. In order to obtain a contradiction from this we need to show we can control the non-linearity of $f^{q_n}|T_n$. From the previous section it follows that this

https://doi.org/10.1017/S0143385700005733 Published online by Cambridge University Press

can be done if the intervals $T_n, f(T_n), \ldots, f^{q_n}(T_n)$ are more or less disjoint. This disjointness will follow from the fact that f preserves ordering on S^1. The proofs of the results in this section can be found in section V of [**He**].

In fact, because f is order-preserving and has no periodic points, there exists an order-preserving and continuous map $h: S^1 \to S^1$ such that $h \circ f = R_\rho \circ h$, where $\rho = \rho(f)$ and R_ρ is the rigid rotation over ρ (see [**He**]). In particular, the order of points in an orbit of f is the same as the order of points in an orbit of R_ρ. Since we keep ρ fixed in this section, we write $R = R_\rho$. Take some $x \in S^1$ and inductively define $q(n)$ to be the time of the nth closest approach:

$$q(0) = 1;$$

$$q(n+1) = \min\{i \in \mathbb{N};\ i > q(n),\ d(x, R^i(x)) < d(x, R^{q(n)}(x))\}.$$

Remark. These numbers $q(n)$ are independent of x (they only depend on $\rho(f)$).

Proposition 4.1. $d(I, f^{q(n)}(I)) \to 0$ as $|n| \to \infty$;

In the next two Propositions we will describe the way the sequence $f^{q(n)}(I)$ is ordered in S^1.

Proposition 4.2. *The intervals* $\{f^{q(2n)}(I);\ n \geq 0\}$ *approach* I *from the right and* $\{f^{q(2n+1)}(I);\ n \geq 0\}$ *approach* I *from the left* (*or vice versa*).

From Propositions 4.1 and 4.2 we know that the sequence $\{f^{q(2n)}(I)\}$, $n \geq 0$, accumulates arbitrarily close to one side of I. In the next Proposition the order of the intervals $\{f^i(I);\ q(2n) \leq i \leq q(2n+2)\}$ is described.

Let $Q(2n)$ be the set of integers $t \in \mathbb{N}$, such that $q(2n) \leq t \leq q(2n+2)$ and such that $f^t(I)$ is between $f^{q(2n+2)}(I)$ and $f^{q(2n)}(I)$.

Proposition 4.3. $t \in Q(2n)$ *if and only if there exists* $i \in \mathbb{N}$ *such that* t *is of the form* $t = q(2n) + iq(2n+1)$ *and* $q(2n) \leq t \leq q(2n+2)$. *In particular, there exists* $a(2n+2) \in \mathbb{N}$ *such that* $q(2n+2) = q(2n) + a(2n+2) \cdot q(2n+1)$.

Proposition 4.4. *Let* T_n *be the smallest interval containing* $f^{-q(n)}(I)$, I *and* $f^{q(n)}(I)$. *Then every point in* S^1 *is an at most three intervals of the sequence*

$$T_n, f(T_n), \ldots, f^{q(n+1)-1}(T_n).$$

5. *Conclusion of the proof of Theorem C for maps* $f \in \mathcal{A}$

Proposition. 5.1. *Let* $f \in \mathcal{A}$ *and assume that* I *is a wandering interval of* f. *There exists* $n_0 < \infty$ *such that if* $n > n_0$ *then*

$$|f^{q(2n)+a(2n+2)\cdot q(2n+1)}(I)| = |f^{q(2n+2)}(I)| > |f^{q(2n+2)-q(2n+1)}(I)|. \tag{5.1}$$

Proof. Let T_n be the smallest interval containing $f^{-q(2n+2)}(I)$, I and $f^{-q(2n+1)}(I)$. Since $T_n \subset [f^{q(2n+1)}(I),\ f^{-q(2n+1)}(I)]$ one has from Proposition 4.4 that each point of S^1 is contained in at most three of the intervals T_n, $f(T_n)$, $f^2(T_n), \ldots, f^{q(2n+2)-1}(T_n)$. Moreover from Proposition 4.1 the length of the two components L_n and R_n of $T_n \backslash I$ tends to zero as $n \to \infty$. Hence we can apply Theorem

3.8 and get that there exist $B_0>0$ and $n_0<\infty$ such that for all $n\ge n_0$, $B(f^{q(2n+2)}, T_n, I)\ge B_0$. Since $f^{q(2n+2)}(R_n)\supset f^{q(2n+2)-q(2n+1)}(I)$ one gets

$$\frac{|f^{q(2n+2)-q(2n+1)}(I)|}{|f^{q(2n+2)}(I)|}\le\frac{|f^{q(2n+2)}(R_n)|}{|f^{q(2n+2)}(I)|}\le\frac{1}{B_0\cdot|I|^2}|L_n||R_n|\frac{|f^{q(2n+2)}(T_n)|}{|f^{q(2n+2)}(L_n)|}\le\frac{|L_n||R_n|}{B_0\cdot|I|^3}.$$

For large n the last term tends to 0 and since $q(2n+2)=q(2n)+a(2n+2)\cdot q(2n+1)$, the Proposition follows. □

PROPOSITION 5.2. *Let $f\in\mathcal{A}$ and I be a wandering interval of f. Then there exists sequences $\{n_k\}$ and $\{i_{n_k}\}$ with $0<i_{n_k}<a(2n_k+2)$ such that for $i=i_{n_k}$,*

$$|f^{q(2n_k)+i\cdot q(2n_k+1)}(I)|<|f^{q(2n_k)+(i-1)\cdot q(2n_k+1)}(I)|, |f^{q(2n_k)+(i+1)\cdot q(2n_k+1)}(I)|. \tag{5.2}$$

(In particular, if $f\in\mathcal{A}$ has a wandering interval then there are infinitely many integers $n\ge 0$ such that $a(2n)>1$.)

Proof. For $n\ge n_0$ (see Proposition 5.1) we have the following. Suppose that for some $n>n_0$ we cannot find an integer $0<i<a(2n+2)$ such that (5.2) is satisfied. Then eq. (5.1) implies that for all $0<i\le a(2n+2)$ one has

$$|f^{q(2n)+i\cdot q(2n+1)}(I)|>|f^{q(2n)+(i-1)\cdot q(2n+1)}(I)|.$$

Hence, because $q(2n+2)=q(2n)+a(2n+2)\cdot q(2n+1)$,

$$|f^{q(2n+2)}(I)|>|f^{q(2n)}(I)|. \tag{5.3}$$

But since $f^i(I)$ are mutually disjoint, $\sum_i |f^i(I)|\le 1$. So there must exist a sequence $n_k\to\infty$ such that (5.3) does not hold and for each of these n_k there exists $0<i<a(2n_k+2)$ such that (5.2) holds. □

Now we can finish the proof of Theorem C.

The proof of Theorem C for maps $f\in\mathcal{A}$. Let n_k and i_{n_k} be as in Proposition 5.2, $q(n)=q(2n_k)+i_{n_k}\cdot q(2n_k+1)$ and K_n the smallest interval containing $f^{q(2n_k+1)}(I)$, I and $f^{-q(2n_k+1)}(I)$. As before let L_n, R_n be the components of $K_n\backslash I$. Proposition 4.4 guarantees the disjointness properties needed in order to apply Theorem 3.8 for $f^{q(n)}|K_n$. From eq. (5.2) and since $B(f^{q(n)}, K_n, I)\ge B_0$,

$$3\cdot\frac{|L_n||R_n|}{|I|^2}\ge\frac{|L_n||R_n|}{|I|^2}\cdot\left[\frac{|f^{q(n)}(I)|}{|f^{q(n)}(R_n)|}+\frac{|f^{q(n)}(I)|}{|f^{q(n)}(R_n)|}\frac{|f^{q(n)}(I)|}{|f^{q(n)}(L_n)|}+\frac{|f^{q(n)}(I)|}{|f^{q(n)}(L_n)|}\right]$$
$$\ge B(f^{q(n)}, K_n, I)\ge B_0.$$

Since $|L_n|$, $|R_n|\to 0$ as $k\to\infty$ we get a contradiction. □

Remark. The difference with the case Yoccoz considers (where there are no intervals where f is constant) is that he can use the operator B_0. This allows him to choose L_n and R_n to be intervals in the backward orbit of I. In this way he can compare $|f^{q(n)}(L_n)|$ and $|R_n|$. This makes his proof easier.

6. *The proof of Theorem C for maps in $\mathcal{A}'$*

The result in this section is a version of a result from unpublished work of C. Gutierrez, A. Lins and W. de Melo.

Assume that $f\in\mathcal{A}'$. This means that f has degree one, is everywhere increasing except on one interval K_f where it is constant and $f|(S^1-K_f)$ is C^1. Moreover outside every neighbourhood of K_f is $\log Df$ of bounded variation, at the boundary points of $K_f = [a, b]$ the (one-sided) derivative of f is infinite and there exists $\varepsilon > 0$ such that $\log Df|[a-\varepsilon, a)$ and $\log Df|(b, b+\varepsilon]$ are monotone.

Let $\Lambda = S^1 \setminus \bigcup_{i\geq 0} f^{-i}(K_f)$.

THEOREM 6.1. *Let $f \in \mathscr{A}'$. Assume that the rotation number of f is irrational. Then there exists $N \in \mathbb{N}$ such that for all $x \in \Lambda$,*

$$Df^N(x) > 1.$$

This Theorem implies Theorem C for maps in $\mathscr{A}'$:

COROLLARY. *$f \in \mathscr{A}'$ (with irrational rotation number) cannot have wandering intervals.*

Proof. This immediately follows from the expansion of $f^N|\Lambda$ and the disjointness of wandering intervals. □

The most important step in proving Theorem 6.1 is the following Proposition.

PROPOSITION 6.2. *Let $f \in \mathscr{A}'$. Assume that the rotation number of f is irrational. Then for each $x \in \Lambda$ there exists $N \in \mathbb{N}$ such that*

$$Df^N(x) > 1.$$

Proof. Choose $x \in \Lambda$, i.e., $f^i(x) \notin K_f$ for all $i \geq 0$. Let $\alpha = -\rho(f)$. We first consider the rotation $R_\alpha : S^1 \to S^1$. From [**He**] (V.8.4) we get that for each $y \in S^1$ and each $n \in \mathbb{N}$,

$$\bigcup_{j=0}^{q_{n+1}-1} R_\alpha^j([y, R_\alpha^{q_n}(y)]) \cup \bigcup_{j=0}^{q_n-1} R_\alpha^j([y, R_\alpha^{q_{n+1}}(y)]) = S^1, \tag{6.1}$$

where the interiors of the above intervals are pairwise disjoint. Using the semi-conjugacy between f^{-1} and R_α and applying (6.1) to f^{-1} we get that x is contained in one of the two unions corresponding to those in (6.1). In particular for all $k \in \mathbb{N}$, there exist $n \geq k$ and $j \in \mathbb{N}$ such that

$$x \in f^{-j}(U), \quad \{\text{int}(U), \text{int}(f(U)), \ldots, \text{int}(f^{-j}(U))\} \text{ are pairwise disjoint}, \tag{6.2}$$

where U is the smallest interval containing K_f and $f^{-q_n}(K_f)$. Because $x \notin \bigcup_{i\geq 0} f^{-i}(K_f)$ one has (for fixed x) that $j \to \infty$ as $k \to \infty$. Let $\varepsilon > 0$ be the number from the definition of the class of maps $\mathscr{A}'$ and let U be a ε neighbourhood of K_f. Let

$$V = \text{Var}(\log Df|U^c), \; A = \max(\log Df|U^c), \quad B = \min(\log Df|U^c).$$

We may assume that $Df|(U\setminus K_f) > 1$ (and therefore $\log Df|(U \setminus K_f) > 0$). Furthermore since $K_f, f^{-1}(K_f), f^{-2}(K_f), \ldots$ is disjoint we may assume that $k \in \mathbb{N}$ is so big that

$$\log \frac{|K_f|}{|f^{-j}(K_f)|} > V + 2A + 3|B|.$$

Now choose $y \in f^{-j}(K_f)$ such that $Df^j(y) = |K_f|/|f^{-j}(K_f)|$. Let T be the smallest interval containing x and y. Notice that $x \in f^{-j}(U)$ and since $U \supset K_f$ and $y \in f^{-j}(K_f)$ one has $T \subset f^{-j}(U)$. Moreover, since $x \notin \bigcup_{i\geq 0} f^{-i}(K_f)$ the interval T is a proper subset of $f^{-j}(U)$. From (6.2) this gives that

$$\{T, f(T), \ldots, f^j(T)\} \text{ is pairwise disjoint}. \tag{6.3}$$

This disjointness and the fact that f is orientation-preserving then implies that the points $\{x, f(x), \ldots, f^j(x)\}$ and $\{y, f(y), \ldots, f^j(y)\}$ are alternatively distributed over S^1 (more precisely, between each two points of the first set there is a point of the second set and vice versa). Choose the usual covering $\mathbb{R} \to S^1 = \mathbb{R} \bmod 1$, $t \to t \bmod 1$, and for a point $z \in S^1$ let $\bar{z}$ be one of the points in $\mathbb{R}$ covering $z \in S^1$. Write the sets

$$\{f^i(x);\, i=0,\ldots,j\} = \{x_i;\, i=0,\ldots,j\},$$
$$\{f^i(y);\, i=0,\ldots,j\} = \{y_i;\, i=0,\ldots,j\},$$

in such a way that there are points $\bar{x}_i$, $\bar{y}_i$, $\bar{a}$, $\bar{b}$ covering x_i, y_i, a, b such that

$$\bar{a} < y_0 = f^j(y) < \bar{b} < \bar{x}_0 < \bar{y}_1 < \bar{x}_1 < \bar{y}_2 < \cdots < \bar{y}_j < \bar{x}_j < \bar{a} + 1.$$

Choose $0 \le s \le r \le j$ such that

$$\bar{y}_s < \bar{b} + \varepsilon < \bar{y}_{s+1} < \cdots \bar{y}_r \le \bar{a} + 1 - \varepsilon < \bar{y}_{r+1}.$$

Let $N = j+1$.

$$\begin{aligned}
\log Df^N(x) &= \sum_{i=0}^{j} \log Df(x_i) \\
&> \sum_{i=0}^{j} \log Df(x_i) - \log Df^j(y) + V + 2A + 3|B| \\
&= \sum_{i=0}^{j} \log Df(x_i) - \sum_{i=1}^{j} \log Df(y_i) + V + 2A + 3|B|.
\end{aligned}$$

Note that there are $j+1$ of the form $\log Df(x_i)$ and only j terms of the form $\log Df(y_i)$. Now we will rewrite (6.6) by pairing the term $\log Df(x_i)$ either with $\log Df(y_i)$ or with $\log Df(y_{i+1})$. Indeed, notice that

$$\begin{aligned}
\bar{a} < \bar{y}_0 = f^j(\bar{y}) < \bar{b} < \bar{x}_0 < \bar{y}_1 < \bar{x}_1 < \bar{y}_2 < \cdots < \bar{x}_{s-1} < \bar{y}_s < b + \varepsilon \\
< \bar{y}_{s+1} < \bar{x}_{s+1} < \bar{y}_{s+2} < \cdots \\
< \bar{x}_{r-2} < \bar{y}_{r-1} < \bar{x}_{r-1} < \bar{y}_r \le a + 1 - \varepsilon \\
< \bar{x}_r < \bar{y}_{r+1} < \bar{x}_{r+1} < \cdots < \bar{y}_j < \bar{x}_j < \bar{a} + 1.
\end{aligned} \tag{6.5}$$

Using that $\log Df$ is decreasing on $(b, b+\varepsilon]$ and increasing on $[a-\varepsilon, a)$ and using (6.4), one gets

$$\begin{aligned}
\log Df^N(x) &> \sum_{i=0}^{j} \log Df(x_i) - \sum_{i=1}^{j} \log Df(y_i) + V + 2A + 3|B| \\
&= \sum_{i=0}^{s-1} [\log Df(x_i) - \log Df(y_{i+1})] + \log Df(x_s) - \log Df(y_{s+1}) \\
&\quad + \sum_{i=s+1}^{r-2} [\log Df(x_i) - \log Df(y_{i+1})] + \log Df(x_{r-1}) - \log Df(y_r) \\
&\quad + \log Df(x_r) + \sum_{i=r+1}^{j} [\log Df(x_i) - \log Df(y_i)] + V + 2A + 3|B| \\
&\ge 0 + \log Df(x_s) - \log Df(y_{s+1}) + \sum_{i=s+1}^{r-2} [\log Df(x_i) - \log Df(y_{i+1})] \\
&\quad \times \log Df(x_{r-1}) - \log Df(y_r) + \log Df(x_r) + 0 + V + 2A + 3|B|.
\end{aligned} \tag{6.6}$$

Now, distinguishing between $\bar{x}_s < \bar{b} + \varepsilon$ or $\bar{x}_s \geq \bar{b} + \varepsilon$, one has

$$\log Df(x_s) - \log Df(y_{s+1}) \geq \min(0 - A, B - A) \geq -|B| - A \tag{6.7a}$$

and similarly

$$\log Df(x_{r-1}) - \log Df(y_r) + \log Df(x_r) \geq \min(B - A + B, B - A + 0) \geq -2|B| - A. \tag{6.7b}$$

Using (6.7) in (6.6) and the definition of V gives

$$\log Df^N(x) \geq 0 + (-|B| - A) - V + (-2|B| - A) + V + 2A + 3|B| > 0.$$

This finishes the proof of the Proposition.

Proof of Theorem 6.1. Let $\Lambda' = S^1 \setminus \bigcup_{i\geq 0} f^{-1} \operatorname{int}(K_f)$. We claim that for all $x \in \Lambda'$ there exists an interval U of x and a number $N \in \mathbb{N}$ such that $Df^N(y) > 1$ for all $y \in U \cap \Lambda$. Indeed if $x \in \Lambda$ this follows from Proposition 6.2, since f^N is C^1 in a neighbourhood of x. If $x \in \Lambda' \setminus \Lambda$ then there exists $n \geq 0$ such that $f^n(x) \in \partial K_f$. Let $N = n + 1$. Since $Df^N(y) \to \infty$ as $y \in \Lambda$ and $y \to x$ the claim also holds for $x \in \Lambda' \setminus \Lambda$. Since Λ' is compact one can cover this set with a finite number of these neighbourhoods U_i as above such that for each i there exists N_i such that $Df^{N_i} | (U_i \cap \Lambda) > 1$. Let

$$N' = \max_i \{N_i\},$$

$$\rho_1 = \min_i \{\min\{\log Df^{N_i} | (U_i \cap \Lambda)\}\} > 0,$$

$$\rho_2 = \min_{0 \leq l \leq N'} \{\min\{\log Df^l | \Lambda\}\}.$$

Since $\rho_1 > 0$ we can choose N so that

$$\rho_1 \cdot \frac{N - N'}{N'} + \rho_2 > 0.$$

Now we will show that $Df^N(x) > 1$ for all $x \in \Lambda$. So choose $x \in \Lambda$. Then there exists a sequence i_l such that $x \in U_{i_1}$, $f^{N_{i_1}}(x) \in U_{i_2}$, $f^{N_{i_1} + N_{i_2}}(x) \in U_{i_3}$, $f^{N_{i_1} + N_{i_2} + N_{i_3}}(x) \in U_{i_4}, \ldots$. Let $k_i(x, N)$ be the number of times U_i appears in the sequence $x, f^{N_{i_1}}(x)$, $f^{N_{i_1} + N_{i_2}}(x), \ldots, f^{N_{i_1} + N_{i_2} + \cdots + N_{i_l}}(x)$ where l is the largest number such that $N_{i_1} + N_{i_2} + \cdots + N_{i_l} \leq N - 1$. By definition of N',

$$N - (N_{i_1} + N_{i_2} + \cdots + N_{i_l}) \leq N'. \tag{6.7}$$

Now write

$$\begin{aligned} \log Df^N(x) &= \log Df^{N_{i_1}}(x) + \cdots + \log Df^{N_{i_l}}(f^{N_{i_1} + N_{i_2} + \cdots + N_{i_{l-1}}}(x)) \\ &\quad + Df^{N - (N_{i_1} + N_{i_2} + \cdots + N_{i_l})}(f^{N_{i_1} + N_{i_2} + \cdots + N_{i_l}}(x)) \\ &\geq \rho_1 \sum_i k_i(x, N) + \log Df^{N - (N_{i_1} + N_{i_2} + \cdots + N_{i_l})}(f^{N - (N_{i_1} + N_{i_2} + \cdots + N_{i_l})}(x) \\ &\geq \rho_1 \sum_i k_i(x, N) + \rho_2 \geq \rho \cdot \frac{N - N'}{N'} + \rho_2 > 1. \end{aligned}$$

Here the second inequality holds because of (6.7) and because of the definition of ρ_2. This finishes the proof of Theorem 6.1. □

7. *The proof of Theorems A and B*

Consider the class $\mathscr{B}$ of C^∞ Cherry vector fields on T^2 such that at each singularity p, $\operatorname{div}(X)(p) \leq 0$. Furthermore, let $\mathscr{B}'$ be the class of C^∞ Cherry vector fields X on T^2 with precisely one hyperbolic sink and one saddle-point and such that X has positive divergence at this saddle-point. We will denote the flow through a point x by $t \to X_t(x)$. Let $\operatorname{Sing}(X)$ be the set of singularities of X.

PROPOSITION 7.1. *Let $X \in \mathscr{B} \cup \mathscr{B}'$. Then there exists a closed C^∞ curve Σ on $T^2 \backslash \operatorname{Sing}(X)$ without self-intersections and with the following properties*

(a) Σ *is everywhere transversal to* X;

(b) Σ *is not retractable to a point.*

Proof. It is enough to show that there exists a recurrent orbit γ which is non-trivial (i.e. not equal to a point or a closed curve), see for example page 144 of [**PaMe**]. Because by assumption f has no periodic orbits, the theorem of Denjoy implies that if X has no singularites then every orbit of X is recurrent and we are done. So assume that X has at least one singularity. Since the Euler characteric of T^2 is zero, this implies that X has at least one saddle-point. Let us call a stable separatrix a component of $W^s(p)\backslash\{p\}$ where p is some saddle-point. We claim that there exists a stable separatrix γ such that the α-limit set of γ, $\alpha(\gamma)$, contains γ. Indeed, take a stable separatrix γ_0. Because of the Denjoy–Schwartz theorem $\alpha(\gamma_0)$ must contain at least one singularity p. Since all singularities of X are sinks or saddles, p is a saddle. Let γ_1 be the stable separatrix of p. Because X has no saddle-connections, $\alpha(\gamma_0) \supset \gamma_1$. By induction we define in this way a sequence of stable separatrices such that $\alpha(\gamma_n) \supset \gamma_{n+1}$ for all $n \geq 0$. Since there are only finitely many separatrices there exist $n \geq 0$, $m > 0$ such that $\gamma_n = \gamma_{n+m}$. One has that

$$\alpha(\gamma_n) \supset \alpha(\gamma_{n+1}) \supset \cdots \supset \alpha(\gamma_{n+m-1}) \supset \gamma_{n+m} = \gamma_n.$$

Hence γ_n is a non-trivial recurrent orbit. As we remarked above this implies the existence of a transversal circle. □

Let $X \in \mathscr{B} \cup \mathscr{B}'$ and let Σ be the closed transversal to X on T^2 from the previous Proposition. Notice that $T^2 \backslash \Sigma$ is an annulus $\Sigma \times (0, 1)$ and we can write $T^2 \cong \Sigma \times [0, 1]/\sim$, where $(s, 0) \sim (s, 1)$. Consider X as a flow on $T^2 \cong \Sigma \times [0, 1]$ where we identify $\Sigma \times \{0\}$ and $\Sigma \times \{1\}$. Since X has no sources it follows that

(c) for every $x \in \Sigma \times \{0\}$ which is not contained in the stable manifold of one of the saddles or one of the sinks, there exists a $t > 0$ such that $X_t(x) \in \Sigma \times \{1\}$;

(d) there exists at least one $x \in \Sigma \times \{0\}$ such that $X_t(x) \in \Sigma \times \{1\}$ for some $t > 0$.

Now let Σ and Σ' be two closed curves transversal to X. Denote the points $x \in \Sigma$ such that there exists $t > 0$ with $X_t(x) \in \Sigma'$ for some $t > 0$ by Σ_0. For $x \in \Sigma_0$, let $t(x)$ be the minimal $t > 0$ such that $X_t(x) \in \Sigma'$ and define the map $f: \Sigma_0 \to \Sigma'$ by $f(x) = X_{t(x)}(x)$. This map is called the *transition map* between Σ and Σ'.

Now let $\Sigma' = \Sigma$ be the section from Proposition 7.1 and take the transition f from Σ to Σ'. This map is called the *return-map* to Σ.

Since orbits of X cannot intersect, f is order preserving. Let I be a component of $(\Sigma \backslash \Sigma_0) \times \{0\}$. Take $x \in \partial I$. Since the basin of a sink is open, x must be contained in the stable manifold of some saddle-point $p \in \Sigma \times (0, 1)$. Since $p \in \partial I$ and X has no saddle-connections, $W^u(p)$ intersects $\Sigma \times \{1\}$ in some point v and for every point in $u \in (\Sigma \backslash I) \times \{0\}$, near ∂I there exists $t > 0$ such that $X_t(u)$ intersects $\Sigma \times \{1\}$ near u. (This can be seen by considering the backward orbits of $X | (\Sigma \times [0, 1])$ of points in $\Sigma \times \{1\}$ near v intersecting $\Sigma \times \{0\}$. This set consists of a neighbourhood of ∂I in Σ.) In particular $\lim_{x \to \partial I, x \notin I} f(x)$ consists of one single point and we can define f on components of $\Sigma \backslash \Sigma_0$ to be constant. From the smooth dependence on initial condition, f is then everywhere continuous, and as smooth as the vector field outside boundary points of $\Sigma \backslash \Sigma_0$. From Theorem 7.1 in the Appendix, f also satisfies the non-flatness conditions from the introduction. It follows that $f \in \mathcal{A}$.

Remark. Let Σ be the section from Proposition 7.1, and $f: \Sigma_0 \to \Sigma$ its transition map. Now $T^2 \backslash \Sigma$ is an annulus A, see figure 7.1. It is easy to see that the orbits through boundary points of Σ_0 go directly to a saddle without leaving the annulus. From this one gets that $\Sigma \backslash \Sigma_0$ consits of finitely many components. Let $I \subset \Sigma$ be one of the components.

(i) I is not equal to a point. Indeed if I consists of one point then this point goes to some saddle-point and both unstable separatrices of this saddle-point intersect Σ. Let $D \subset A$ be the region bounded by Σ and these two unstable separatrices, see figure 7.1. But since X has no periodic orbits and no sources, this contradicts the Poincaré–Bendixon theorem.

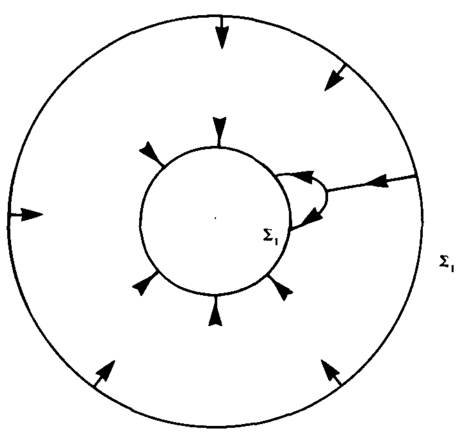

FIGURE 7.1. This situation cannot occur for $X \in \mathcal{B} \cup \mathcal{B}'$.

(ii) Let $I = [a, b]$. If p is the saddle-point such that $a \in W^s(p)$ then also $b \in W^s(p)$. Indeed, let γ be the stable separatrix of p which does not contain a. From the assumption on X and Poincaré–Bendixon theorem it follows again that γ cannot be contained in the annulus. So let c be the intersection of the closure of $\gamma \cap A$ and Σ. It is easy to see that $c = b$. From all this it follows that the situation is as drawn in figure 7.2.

Now we are in the position to prove Theorems A and B.

https://doi.org/10.1017/S0143385700005733 Published online by Cambridge University Press

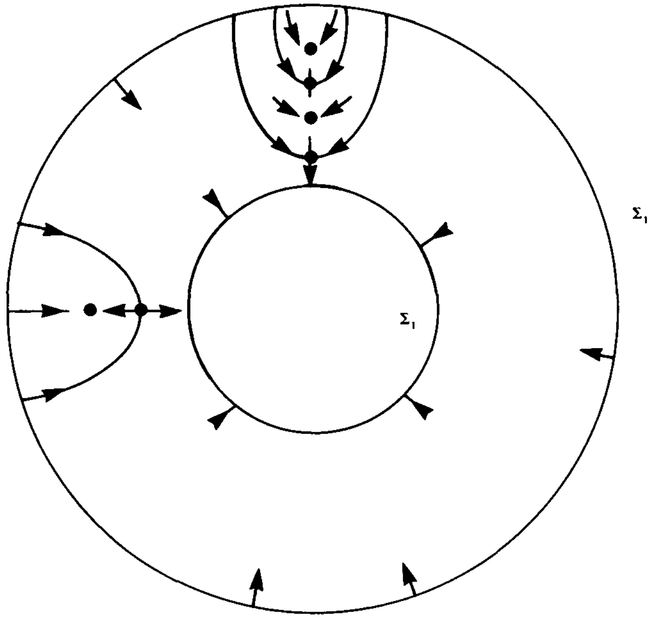

FIGURE 7.2. A typical situation for $x \in \mathcal{B} \cup \mathcal{B}'$.

Proof of Theorem A. Since to a vector field X we associated a map $f \in \mathcal{A}$, and since X has no periodic orbits and has no saddle-connections, $\rho(f)$ is irrational.

But if $\rho(f)$ is irrational, we can apply the Corollary to Theorem C to $f: S^1 \cong \Sigma \to \Sigma \cong S^1$ and the orbit of every point in $\{x; f^n(x) \in \Sigma_0, \forall n \geq 0\}$ is dense in this set. But now notice that every point in $\Sigma \backslash \Sigma_0$ is contained in the basin of a periodic attractor or in the stable manifold of a saddle-point. The statement follows. □

Proof of Theorem B. Take X as above and let Σ be the section from Proposition 7.1. There exists a neighbourhood U of X such that for each $Y \in U$, Σ is also a section of Y. Associate $f_Y: \Sigma \to \Sigma$ to Y as above. From the construction f_Y depends continuously on $Y \in U$. If there exists a conjugacy between $f_X, f_Y: \Sigma \to \Sigma$ then one can construct an equivalence between X and Y in the same way as in the proof of the structural stability of Morse–Smale flows on surfaces, see for example, [**PaMe**].

8. *Appendix: The transition map near singularities*

Let X be a C^∞ vector field on $\mathbb{R}^2$ which has a hyperbolic singularity at 0 of saddle-type with eigenvalues $\lambda > 0 > \mu$. Let p_1 and p_2 be points in $W^s(0)$ respectively $W^u(0)$. Furthermore let Σ_i be a C^2 curve through p_i which is transversal to X. If we choose Σ_1 sufficiently small, then for every x in one of the components of $\Sigma_1 - \{p_1\}$ there exists $t \geq 0$ such that $X_t(x) \in \Sigma_2$. Call this component Σ_1' and let $t(x) \in \mathbb{R}$ be the smallest number so that $X_{t(x)}(x) \in \Sigma_2$ and define $T: \Sigma_1' \to \Sigma_2$ by

$$T(x) = X_{t(x)}(x).$$

We call this map the *transition map.*

Since Σ_i are C^2 curves, it is a well defined notion to say that maps $\phi:\Sigma_1\to\mathbb{R}$ and $\psi:\mathbb{R}\to\Sigma_2$ have a derivative which has bounded variation.

In this section we want to show that $T:\Sigma_1\to\Sigma_2$ is equal to a map ϕ_α, up to maps whose derivatives have bounded variations. Here $\alpha=|\mu|/\lambda>0$, and ϕ_α is defined by $\phi_\alpha=\pm|z|^\alpha$.

THEOREM 8.1. *Let X be a C^∞ vector field on $\mathbb{R}^2$ which has a hyperbolic singularity at* 0 *of saddle-type with eigenvalues $\lambda>0>\mu$. Let $\alpha=|\mu|/\lambda$. There exist maps $\phi,\psi:\mathbb{R}^+\to\mathbb{R}^+$ which are C^1, such that $D\phi$ and $D\psi$ have bounded variation and such that the map T from above is of the form*

$$T(x)=\phi\circ\phi_\alpha\circ\psi(x). \tag{8.1}$$

Proof.

Step 1. Let us first show that (8.1) is true if X is linear near 0.

If Σ_i are the straight lines $\{x=1\}$, $\{y=1\}$ and $T_{\{x=1\},\{y=1\}}$ the transition map from $\{x=1\}$ and $\{y=1\}$, then (8.1) follows from explicit integration (and choosing the natural parametrisations of $\{x=1\}$, $\{y=1\}$. If Σ_i are different C^2 curves (8.1) still holds. In fact consider the transition map T_1 from Σ_1 to $\{x=1\}$, resp. $T_2:\{y=1\}\to\Sigma_2$. Then $T_{\Sigma_1,\Sigma_2}=T_2\circ T_{\{x=1\},\{y=1\}}\circ T_1$. Since X is linear (and in particular the flow map is C^2) it follows from the implicit function theorem that the maps T_i are C^2. This last argument also shows that if (8.1) is true for one choice of C^2 curves Σ_i, then it is also true for any other choice of C^2 curves $\tilde{\Sigma}_i$ as above.

Step 2. X is C^2 linearisable if $|\lambda|\notin\{|\mu|,2|\mu|,\frac{1}{2}|\mu|\}$.

According to the Linearisation Theorem of Sternberg (see [**Ster**], [**Bel**], [**BD**] or for example Theorem 4 in [**Sto**]) there exists a C^2 coordinate system ϕ near 0 such that ϕ_*X is linear if

$$|\lambda|\neq|\mu|,2|\mu|,\tfrac{1}{2}|\mu|. \tag{8.2}$$

That is, in this case

$$\phi_*X(x,y)=\lambda\cdot x\frac{\partial}{\partial x}+\mu\cdot y\frac{\partial}{\partial y}.$$

Using step 1 we are finished if (8.2) holds.

Step 3. A normal form for X when $|\mu|=2|\lambda|$ or $|\mu|=|\lambda|$.

Let us now see what happens if (8.2) fails. We deal with the cases that $|\mu|=2|\lambda|$ and that $|\mu|=|\lambda|$. The case that $|\lambda|=2|\mu|$ is similar. By considering a multiple of the vector field we may assume that X is of the form

$$X(x,y)=x\cdot\frac{\partial}{\partial x}-2y\cdot\frac{\partial}{\partial y}+o|\underline{x}|^2. \tag{8.3a}$$

$$X(x,y)=x\cdot\frac{\partial}{\partial x}-y\cdot\frac{\partial}{\partial y}+o|\underline{x}|^2. \tag{8.3b}$$

Here $\underline{x}=(x,y)$.

LEMMA 8.2. *Let X be a C^∞ vector field.*

(a) *If X is of the form* (8.3a), *then there exists a C^∞ change of coordinates ϕ such that*

$$\phi_*(X) = x(1+ax^2y)\frac{\partial}{\partial x} - 2y(1+bx^2y)\frac{\partial}{\partial y} + o|\underline{x}|^6. \tag{8.4a}$$

(b) *If X is of the form* (8.3b), *then there exists a C^∞ change of coordinates ϕ such that*

$$\phi_*(X) = x(1+axy)\frac{\partial}{\partial x} - y(1+bxy)\frac{\partial}{\partial y} + o|\underline{x}|^4. \tag{8.4b}$$

Proof. Let X_1 be the vector field on $\mathbb{R}^2$ which has the same 1-jet in 0 as X and whose coefficients are linear. Let $[X_1, -]_h : H^h \to H^h$ be the linear map which assigns to each homogeneous vector field $Y \in H^h$ of degree h the Lie-product $[X_1, Y]$ (which is again in H^h). For each $h \in \mathbb{N}$ choose G^h so that $H^h = \mathrm{Im}([X_1, -]_h) + G^h$. According to Theorem 2.1 of Takens [**Ta**] there exists for each $l < \infty$, a C^∞ diffeomorphism Φ such that $\Phi_*(X)$ is of the form

$$\Phi_*(X) = X_1 + g_2 + \cdots + g_l + o|\underline{x}|^l)$$

where $g_i \in G^i$, $i = 2, 3, \ldots, l$.

Let us determine $\mathrm{Im}([X_1, -]_h)$. Taking $\beta \in \{1, 2\}$,

$$\left[x\frac{\partial}{\partial x} - \beta y\frac{\partial}{\partial y}, x^n y^m \frac{\partial}{\partial x}\right] = (n - \beta m - 1)x^n y^m \frac{\partial}{\partial x},$$

$$\left[x\frac{\partial}{\partial x} - \beta y\frac{\partial}{\partial y}, x^n y^m \frac{\partial}{\partial y}\right] = (n - \beta m + \beta)x^n y^m \frac{\partial}{\partial y}.$$

For $\beta = 2$, we can choose $G^h = 0$ for $h \leq 6$, $h \neq 4$, and

$$G^4 = \left\langle x^3 \cdot y \cdot \frac{\partial}{\partial x}, x^2 \cdot y^2 \cdot \frac{\partial}{\partial y}\right\rangle.$$

Case (a) follows.

For $\beta = 1$, we can choose $G^h = 0$ for $h \leq 4$, $h \neq 3$, and

$$G^3 = \left\langle x^2 \cdot y \cdot \frac{\partial}{\partial x}, x \cdot y^2 \cdot \frac{\partial}{\partial y}\right\rangle.$$

Case (b) follows and the proof of the Lemma is finished. □

Step 4. A C^1 + 'bounded variation' linearisation of X.

From step 1, we may assume that X is of the form

$$X = x(1 + a \cdot x^2 y)\frac{\partial}{\partial x} - 2y(1 + b \cdot x^2 y)\frac{\partial}{\partial y} + o|\underline{x}|^6; \tag{8.5a}$$

or

$$X = x(1 + a \cdot x \cdot y)\frac{\partial}{\partial x} - y(1 + b \cdot x \cdot y)\frac{\partial}{\partial y} + o|\underline{x}|^4. \tag{8.5b}$$

Step 4A. Let us first consider the case of (8.5a). Consider the C^1 coordinate transformation $\Psi(x, y) = (\tilde{x}, \tilde{y})$ defined by:

$$\tilde{x} = x - ax^3 y \log|x|,$$
$$\tilde{y} = y - bx^2 y^2 \log|y|.$$

Then $\{x=1\} = \{\tilde{x}=1\}$ and $\{y=1\}=\{\tilde{y}=1\}$. Moreover, the transformation between y and $\tilde{y}$ coordinates on the line $\{x=1\}=\{\tilde{x}=1\}$, $\{(x,y); x=1\} \to \{(\tilde{x},\tilde{y}); \tilde{x}=1\}$, is equal to $y \mapsto \tilde{y} = y - by^2 \log|y|$, which has a derivative of bounded variation on $[-1, 1]$. Similarly $\tilde{y} \mapsto y$ is also C^1 and has a derivative of bounded variation. The same holds for the coordinate transformation $\{(x,y); y=1\} \to \{(\tilde{x},\tilde{y}); \tilde{y}=1\}$. Since $T_{\{x=1\},\{y=1\}}$ is equal to $T_{\{\tilde{x}=1\},\{\tilde{y}=1\}}$ composed with $\tilde{y} \mapsto y$ on the left and with $x \mapsto \tilde{x}$ on the right, it suffices therefore to prove the result for $\Psi_* X$. Let us calculate $\Psi_* X$.

$$\left(X, \frac{\partial}{\partial \tilde{x}}\right) = \left(X, \frac{\partial}{\partial x}\right) \cdot \frac{\partial \tilde{x}}{\partial x} + \left(X, \frac{\partial}{\partial y}\right) \cdot \frac{\partial \tilde{x}}{\partial y}$$
$$= x(1+ax^2y)(1 - a3x^2 y \log|x| - ax^2 y) - 2y(1+bx^2y)(-ax^3 \log|x|) + o|\underline{x}|^6$$
$$= x - ax^3 y \log|x| + o|\underline{x}|^6 + o|\tilde{\underline{x}}|^6 = \tilde{x} + o|\tilde{\underline{x}}|^6,$$

and

$$\left(X, \frac{\partial}{\partial \tilde{y}}\right) = \left(X, \frac{\partial}{\partial x}\right) \cdot \frac{\partial \tilde{y}}{\partial x} + \left(X, \frac{\partial}{\partial_y}\right) \cdot \frac{\partial \tilde{y}}{\partial y}$$
$$= x(1+ax^2y)(-2bxy^2 \log|y|) - 2y(1+bx^2y)(1 - 2bx^2 y \log|y| - bx^2 y)$$
$$+ o|\underline{x}|^6$$
$$= -2y + 2bx^2y^2 \log|y| + o(|\underline{x}|^6) + o|\tilde{\underline{x}}|^6 = -2\tilde{y} + o|\tilde{\underline{x}}|^6.$$

Here $\tilde{\underline{x}} = (\tilde{x}, \tilde{y})$. Hence

$$\Psi_* X = \tilde{x} \cdot \frac{\partial}{\partial \tilde{x}} - 2 \cdot \tilde{y} \cdot \frac{\partial}{\partial \tilde{y}} + o|\tilde{\underline{x}}|^6.$$

Using the notation of Theorem 2 of [**Sto**] we take $\sigma = 2$, $r = 6$, and it follows that there exists a transformation of Hölder class C^s, where $s = \sigma(r-1)/1+\sigma = 10/3$ linearising $\Psi_* X$ at 0. (We use here the vector field analogue of these linearisation results for diffeomorphisms, see the end of the introduction in [**Sto**].) Using step 1, the result follows for (8.5a).

Step 4B. Now we consider the case of (8.5b). Consider the C^1 coordinate transformation $\Psi(x, y) = (\tilde{x}, \tilde{y})$ defined by:

$$\tilde{x} = x - ax^2 y \log|x|,$$
$$\tilde{y} = y - bxy^2 \log|y|.$$

Then as before $\{(x,y); x=1\} = \{(\tilde{x},\tilde{y}); \tilde{x}=1\}$ and $\{(x,y); y=1\} = \{(\tilde{x},\tilde{y}); \tilde{y}=1\}$ and the coordinate transformation $\tilde{y} \mapsto y$ and $x \mapsto \tilde{x}$ on these curves are C^1 and have derivatives which have bounded variation. So as before it suffices to prove the result for $\Psi_* X$. Let us calculate $\Psi_* X$.

$$\left(X, \frac{\partial}{\partial \tilde{x}}\right) = \left(X, \frac{\partial}{\partial x}\right) \cdot \frac{\partial \tilde{x}}{\partial x} + \left(X, \frac{\partial}{\partial y}\right) \cdot \frac{\partial \tilde{x}}{\partial y}$$
$$= x(1+axy)(1-a2xy \log|x| - axy) - y(1+bxy)(-ax^2 \log|x|) + o|\underline{x}|^4$$
$$= x - ax^3 y \log|x| + o|\underline{x}|^4 + o|\tilde{\underline{x}}|^4 = \tilde{x} + o|\tilde{\underline{x}}|^4,$$

and

$$\left(X, \frac{\partial}{\partial \tilde{y}}\right) = \left(X, \frac{\partial}{\partial x}\right) \cdot \frac{\partial \tilde{y}}{\partial x} + \left(X, \frac{\partial}{\partial y}\right) \cdot \frac{\partial \tilde{y}}{\partial y}$$
$$= x(1+axy)(-by^2 \log|y|) - y(1+bxy)(1-2bxy \log|y| - bxy) + o|\underline{x}|^4$$
$$= y + bxy^2 \log|y| + o|\underline{x}|^4) + o|\tilde{\underline{x}}|^4 = \tilde{y} + o|\tilde{\underline{x}}|^4.$$

Hence

$$\Psi_* X = \tilde{x} \cdot \frac{\partial}{\partial \tilde{x}} - \tilde{y} \cdot \frac{\partial}{\partial \tilde{y}} + o|\underline{x}|^4.$$

Using the notation of Theorem 2 of [Sto] we take $\sigma = 1$, $r = 4$ and it follows that there exists a transformation of Hölder class C^s, where $s = (r+\sigma)/(1+\sigma) = \frac{5}{2}$ linearising $\Psi_* X$ at 0. Using step 1, the result follows.

Since we had reduced everything to these two cases we are finished with the proof of Theorem 8.1. □

Remark. All the results in this section are valid if X is C^6.

REFERENCES

[Bel] G. R. Belitskii. Equivalence and normal forms of germs of smooth mappings. *Russian Math. Surveys* **33** (1978), 107–177.

[BD] P. Bonckaert and F. Dumoriter. On a linearization theorem of Sternberg for germs of diffeomorphisms. *Math. Z.* **185** (1984), 115–135.

[Ch] T. M. Cherry. Analytic quasi-periodic curves of discontinuous type on the torus. *Proc. London Math. Soc.* **44** (1938), 175–215.

[Den] A. Denjoy. Sur les courbes définies par les équations différentielles à la surface du tore. *J. de Math. Pures et Appl.* **11** (1932), 333–375.

[Ha] G. Hall. A C^∞ Denjoy counter example. *Ergod. Th. & Dynam. Sys.* **1** (1981), 261–272.

[He] M. R. Herman. Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations. *Publ. Math. I.H.E.S.* **49** (1979), 5–233.

[MS1] W. de Melo and S. van Strien. A structure theorem in one dimensional dynamics. *Ann. Math.* **189** (1989), 519–546.

[MS2] W. de Melo and S. van Strien. One-dimensional dynamics: The Schwarzian derivative and beyond. *Bull. Amer. Math. Soc.* **18** (1988), 159–162.

[PaMe] J. Palis and W. de Melo. *Geometric Theory of Dynamical Systems, an Introduction.* Springer Verlag: New York, 1982.

[Sch] A. J. Schwartz. A generalization of a Poincaré-Bendixson Theorem to closed two dimensional manifolds. *Amer. J. Math.* **85** (1963), 453–458.

[Ster] S. Sternberg. On the structure of local homeomorphisms in Euclidean *n*-space, II. *Amer. J. Math.* **80** (1958), 623–631.

[Sto] D. Stowe. Linearisation in two dimensions. *J. Diff. Eq.* **63** (1986) 183–226.

[Ta] F. Takens. Singularities of vector fields. *Publ. Math. I.H.E.S.* **43** (1974), 47–100.

[Y] J. C. Yoccoz. Il n'y a pas de contre-exemples de Denjoy analytique. *C.R. Acad. Sci. Paris* **298**, ser. 1, no. 7 (1984), 141–144.