Pacific Journal of Mathematics

MINIMIZING THE NUMBER OF FIXED POINTS FOR SELF-MAPS OF COMPACT SURFACES

Michael R. Kelly

MINIMIZING THE NUMBER OF FIXED POINTS
FOR SELF-MAPS OF COMPACT SURFACES

Michael R. Kelly

Let P denote the topological space obtained by taking a closed regular neighborhood of the figure-eight in the plane. Let $MF(f)$ denote the minimum number of fixed points achievable among maps homotopic to a given self-map f of P. We present here a formula for the value of $MF(f)$. Note that $MF(f)$ depends on the induced homomorphism, $f_\#$, on fundamental group, so our formula concerns the two relevant words in the free group on the letters a and b corresponding to the loops which comprise the figure eight. Special case: Let $g_m: P \to P$, $m \geq 0$, be given such that $(g_m)_\#(a) = (bab^{-1}a^{-1})^m ba$ and $(g_m)_\#(b) = 1$. It is easy to show that the Nielsen number of g_m, $N(g_m)$, is equal to zero. On the other hand, our formula shows that $MF(g_m) = 2m$. Hence the difference between $N(f)$ and $MF(f)$ can be made arbitrarily large.

Introduction. Let X^n be a compact n-dimensional manifold (with or without boundary) and let f be a self-map of X. We are interested in the problem of finding $MF(f) \equiv$ minimum number of fixed points occurring among maps homotopic to f. An imporatnt invariant which arises in this problem is the Nielsen number $N(f)$, defined as the number of essential fixed point classes of f. As the precise definition does not play a role in the results of this paper the reader is referred to [B] or [J1] for an exposition.

One reason for the importance of the Nielsen number is the inequality $N(f) \leq MF(f)$; even when X is a compact ANR. Another is the following classical result.

THEOREM 0. *If $n > 2$ or if X is a surface with non-negative Euler characteristic then $MF(f) = N(f)$.*

In the case $n > 2$ this was first proved by Wecken [W]. A modern treatment, using the Whitney Lemma, is given in [J4]. For surfaces with $\chi(X) \geq 0$ the reader is referred to the work of Nielsen [N] using the torus. Also, as a result of Nielsen's and Thurston's work with hyperbolic surfaces, $MF(f) = N(f)$ when f is a homeomorphism (see [BK] or [J5]). On the other hand, Jiang [J2], [J3] has produced examples for which $N(f) = 0$ but $MF(f) > 0$.

In this paper we give a formula for the value of $MF(f)$ for any self-map of the disk with two open holes removed (see Theorems 1.1 and

1.4). As a consequence, in Corollary 1.2, we show that the difference between $MF(f)$ and $N(f)$ can be arbitrarily large. The methods used in proving Theorem 1.1 are quite different than the braid theoretic techniques introduced by Fadell and Husseini [**FH**] and employed by Jiang in his examples. Motivated by the "spurious elements" in [**FH**; §5], we judiciously apply homotopies so as to find a representative map g, homotopic to a given f, having $MF(f)$ fixed points.

I have been informed that X. G. Zhang has also shown that the difference between $MF(f)$ and $N(f)$ can be arbitrarily large.

This paper was submitted as partial fulfillment of the requirements for the Ph.D. at the State University of New York at Binghamton under the supervision of Professor Ross Geoghegan.

1. The main result. Let S_1 and S_2 denote simple closed curves in $\mathbf{R}^2$ meeting in the point x_0. Set $P_0 = S_1 \cup S_2$ and let P denote the topological space obtained by taking a closed regular neighborhood of P_0 in $\mathbf{R}^2$ (i.e., $P \equiv$ pair of pants). Orient S_1 and S_2 so that one of the boundary components of P is freely homotopic to the loop obtained by traversing first S_1 then S_2 (following orientation). We identify $\pi_1(P, x_0)$ with the free group, G, on the letters a, b by associating the homotopy classes of the oriented loops S_1 and S_2 with a and b respectively. See Figure 1 for an illustration.

Let $r: P \to P_0$ be a retraction such that $r^{-1}(x)$ is an arc for each point $x \neq x_0$ and $r(A_i)$, $1 \leq i \leq 4$, is a single point (see Figure 1). Given reduced words $X, Y \in G$ we define $f_{(X,Y)}: P_0 \to P_0$ by sending S_1 and S_2 to the loops corresponding to the words X and Y respectively. In doing so, $P_0 \cap P_C$ together with a finite set of points in $P_0 - (P_C \cap P_0)$ are mapped to x_0 while each complimentary domain is mapped homeomorphically onto a component of $P_0 - \{x_0\}$. Define $F_{(X,Y)}: P \to P$ by $F_{(X,Y)} = f_{(X,Y)} \circ r$. These self-maps of P will act as "reference-maps" for all future work. Note that $(F_{(X,Y)})_\#(a) = X$, $(F_{(X,Y)})_\#(b) = Y$, and, given a pair (X, Y), it is easy to calculate the number of fixed points for $F_{(X,Y)}$.

Convention Unless stated otherwise, it is assumed that all elements of G and their products are reduced words. For example, $X = W\overline{X}$ means that W, $\overline{X}$ are reduced and $(W_e)^{-1} \neq \overline{X}_b$ (see (2) below).

Notation (1) Following standard conventions; $\mathrm{Fix}(f) = \{x \mid f(x) = x\}$, ' $\simeq$ ' means 'homotopic to', ∂M denotes the manifold boundary of the manifold M and $\mathring{M}$ its interior.
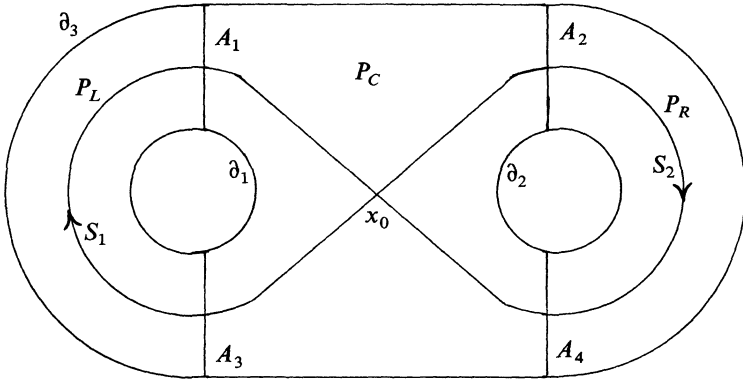
FIGURE 1. The pair of pants—illustrating $P_0 = S_1 \cup S_2$,
$P = P_L \cup P_C \cup P_R$, $A = A_1 \cup A_2 \cup A_3 \cup A_4$, $\partial P = \partial_1 \cup \partial_2 \cup \partial_3$

(2) Given $W \in G$ let $|W|$ denote the length of W. If $W \neq 1$, let W_b denote the prefix and W_e the suffix of W with $|W_b| = |W_e| = 1$. Also, $1_b = 1, 1_e = 1, W_b^{-1} = (W_b)^{-1}$, and $W_e^{-1} = (W_e)^{-1}$.

(3) If $\mathcal{W}$ is a finite set of words and X is a single word in G define $\Phi_{\mathcal{W}}(X) = \sum_{W \in \mathcal{W} \cup \mathcal{W}^{-1}}$ (number of appearances of W in X) where $\mathcal{W}^{-1} = \{ R^{-1} | R \in \mathcal{W} \}$.

(4) If l is a letter and W, X, Y are words in G let

$$\rho_l(W) = \begin{cases} 2 & \text{if } W_b = l = W_e \text{ and } |W| > 1 \\ 0 & \text{if } W_b \neq l \text{ and } W_e \neq l \\ 1 & \text{otherwise} \end{cases}$$

and

$$\sigma(X, Y) = \begin{cases} -3 & \text{if } \rho = 4 \\ -2 & \text{if } \rho = 3 \\ -1 & \text{if } \rho = 1 \text{ or } 2 \\ +1 & \text{if } \rho = 0 \end{cases} \quad \text{where } \rho = \rho_a(X) + \rho_b(Y).$$

Let $X, Y \in G$ be given and consider the following conditions on the ordered pair (X, Y):

Condition (T_1): $Y = 1$ and X is cyclically reduced.

Condition (T_2): $Y \neq 1$ is cyclically reduced and $X = U\bar{X}U^{-1}$ where $\bar{X} \neq 1$ is cyclically reduced. $U = 1$ implies that $X_b \neq Y_b$, $X_b \neq Y_e^{-1}$, $X_e \neq Y_e$, and $X_e \neq Y_b^{-1}$.

Condition (T_3): $X = W\bar{X}$, $Y = W\bar{Y}$ where $W \neq 1$ is the maximal common prefix of X and Y. Also, $\bar{X} = 1$ implies that $\bar{Y} = 1$, and $X_e = Y_e$ implies that $\bar{Y} = 1$ and there exists V such that $X = V^r, Y = V^s, r \geq s > 0$.

Condition (T_4): $X = W\overline{X}$, $Y = \overline{Y}W^{-1}$ where $W \neq 1$ is the maximal condition prefix of X and Y^{-1}. Also, $\overline{X} \neq 1$, $\overline{Y} \neq 1$, and $X_e \neq Y_b^{-1}$.

Condition (T_5): $X = l_1 l_2 \overline{X}, Y = l_1^{-1}$ where $l_1 l_2 \in \{ ab, ba, a^{-1}b^{-1}, b^{-1}a^{-1}\}$ and $\overline{X} \neq 1$ implies that $\overline{X}_e = l_2$.

The ordered pair (X, Y) *is of type* T_K, $1 \leq K \leq 5$, iff (X, Y) satisfies condition (T_K) but does not satisfy condition $(T_{K'})$ when $K' \neq K$. Notice that the conditions are not mutually disjoint. For example, (aba^{-2}, ab^{-1}) satisfies both T_2 and T_3. But, up to free homotopy of maps, this pair is the same as $(ba^{-1}, b^{-1}a)$ which is of type T_2. This idea is developed further in Theorem 1.4.

We define a function M from $\{(X, Y) | (X, Y)$ has a type$\}$ into the nonnegative integers as follows. If (X, Y) is of type T_K then.

For $K = 1$; $M(X, Y) = \Phi_a(X) + \lambda_1$ where

$$\lambda_1 = \begin{cases} -1 & \text{if } a \text{ appears in } X \\ +1 & \text{if } a \text{ does not appear in } X. \end{cases}$$

For $K = 2$; If $U_e \overline{X}_b = aa$ or $\overline{X}_e U_e^{-1} = aa$ or $\rho_a(\overline{X}) = 2$ or $\Phi_{aa}(U) > 0$ then

$$M(X, Y) = \Phi_a(\overline{X}) + \Phi_b(Y) + \Phi_{\{ab, ab^{-1}, a^{-1}b, a^{-1}b^{-1}\}}(U)$$

$$+ \Phi_a(U_b) + \sigma(X, Y) + \lambda_2$$

where

$$\lambda_2 = \begin{cases} -1 & \text{if } U_e \overline{X}_b = aa \text{ or } \overline{X}_e U_e^{-1} = aa \text{ or } \rho_a(\overline{X}) = 2 \\ 0 & \text{if } U_e = b^{\pm 1} \text{ and } \rho_a(\overline{X}) < 2 \\ +1 & \text{otherwise} \end{cases}$$

otherwise $M(X, Y) = \Phi_a(X) + \Phi_b(Y) + \sigma(X, Y)$.

For $K = 3$; If $\overline{Y} = 1$ write $X = W^{N-1}\overline{W}\overline{X}$ where N is chosen maximal and $\overline{W}$ is the maximal common prefix of W and the reduced form of $W^{1-N}\overline{X}$. Then $M(X, Y) = \Phi_a(X) + \Phi_b(Y) + \min\{\lambda_3, \sigma(X, Y)\}$ where

$$\lambda_3 = \begin{cases} -2 & \text{if either } \overline{X}_b \overline{Y}_b = ab \text{ or } \left(\overline{Y} = 1 \text{ and } \overline{\overline{X}}_b(\overline{W}^{-1}W)_b = ab\right). \\ -1 & \text{if } \lambda_3 = -2 \text{ is not satisfied and if either } W_e \overline{X}_b = b^{-1}a \text{ or } \\ & W_e \overline{Y}_b = a^{-1}b \text{ or } \Phi_{a^{-1}b}(W) > 0 \text{ or} \\ & \left(\overline{Y} = 1 \text{ and } (W^{N-1}\overline{W})_e \overline{\overline{X}}_b = b^{-1}a\right). \\ +1 & \text{otherwise.} \end{cases}$$

For $K = 4$; $M(X, Y) = \Phi_a(X) + \Phi_b(Y) + \tau$ where the value of τ is obtained from the following table.

	CONDITION	VALUE OF τ
(I)	$\overline{X} = aX'$, $\overline{Y} = Y'b$ or $\overline{X} = aX'$, $W_e = b$ (set $Y' = Y$) or $\overline{Y} = Y'b$, $W_e = a^{-1}$ (set $X' = X$)	$\lambda_4 - \Phi_{\{ab, ba\}}(W) - 2$
(II)	not(I) but $W = W_1(ba)^\delta W_2$ where $\delta = \pm 1$ and $\Phi_{ba}(W_2) = 0$ (set $X' = X, Y' = Y$)	$\lambda_4 - \Phi_{\{ab, ba\}}(W_1(ba)^\delta) - 1$
(III)	neither (I) nor (II)	$\sigma(X, Y)$

Here

$$\lambda_4 = \begin{cases} +1 & \text{if } X'_e \neq a, Y'_b \neq b, \text{ and } W_b = a^{-1} \text{ or } b. \\ 0 & \text{if } W_b = a \text{ or } b^{-1} \text{ and } \rho_a(X'_e) + \rho_b(Y'_b) < 2. \\ -1 & \text{otherwise.} \end{cases}$$

For $K = 5$; $M(X, Y) = \Phi_a(X) + \Phi_b(Y) + \sigma(X, Y) - 2\rho_a(\overline{X}_e)$.

We are now prepared to state the main result of this paper. This is given in Theorems 1.1 and 1.4. $MF(f) \equiv \text{Min}\{\#\text{Fix}(g) \mid g \simeq f\}$.

THEOREM 1.1 *Let* (X, Y) *have a type. Then* $MF(F_{(X, Y)}) = M(X, Y)$.

COROLLARY 1.2 *For each integer* $m \geq 0$, *let* $W(m) = (bab^{-1}a^{-1})^m ba$ *and* $g_m = F_{(W(m), 1)}$. *Then* $N(g_m) = 0$ *and* $MF(g_m) = 2m$.

Proof. Clearly, $(W(m), 1)$ is of type T_1. As $M(W(m), 1) = 2m$, Theorem 1.1 gives $MF(g_m) = 2m$. That $N(g_m) = 0$ follows easily from the definition of the Nielsen number. $\square$

Let ϕ and ψ be homeomorphisms of P such that $\phi_\#(a) = b$, $\phi_\#(b) = a$, $\psi_\#(a) = ab$, $\psi_\#(b) = a^{-1}$. Given $f: P \to P$ define three *moves*:

Move (I). Choose $Q \in G$, the free group on a and b, and replace f by f' where f' is homotopic to f and $f'_\#(\cdot) = Qf_\#(\cdot)Q^{-1}$ on π_1.

Move (II): Replace f by $\phi f \phi^{-1}$.

Move (III): Replace f by $\psi f \psi^{-1}$.

Procedure 1.3. Given $f_1: P \to P$ and $1 \leq i \leq 5$, f_{i+1} is obtained from f_i by applying Step (i) as follows: let $(R_i, S_i) \equiv ((f_i)_\#(a), (f_i)_\#(b))$ and as a convention for Step (i), if (R_i, S_i) has a type we set $f_{i+1} = f_i$ (hence $(R_{i+1}, S_{i+1}) = (R_i, S_i)$).

Step (1). Apply Move (I) so that (a) $|R_2| + |S_2| \leq |VR_2V^{-1}| + |VS_2V^{-1}|$ for each $V \in G$, (b) S_2 is cyclically reduced, and (c) $(R_2)_e = (S_2)_e$ or $(R_2)_e = (S_2)_b^{-1}$ implies that $R_2 = W^n$ and $S_2 = W^m$ for some $W \in G$ and integers n, m. We show in 1.4 that this step can be carried out.

Step (2). Apply Move (II) iff either (R_3, S_3) has a type or $(R_2, S_2) = (V^{-1}, VS)$ for some $V \in G$.

Step (3). Apply Move (III). (Recall: only if (R_2, S_2) does not have a type.)

Step (4). Apply Move (I) as in Step (1).

Step (5). Apply Move (II).

THEOREM 1.4. *Given* $f: P \to P$, *Procedure* 1.3 *yields a pair* (R, S) *having a type, such that* $MF(f) = MF(F_{(R,S)})$. *Hence* $MF(f) = M(R, S)$.

Proof (of 1.4). Notice that the value of $MF(\cdot)$ is invariant under each of the three moves. We show that the pair (R_6, S_6) in Procedure 1.3 has a type. First, to see that Step (1) can be carried out, note that properties (a) and (b) are easy to achieve. For (c), conjugation by $Q = R_e$ has the effect of shifting $R_e = S_e$ (or $R_e = S_b^{-1}$) to $R_b = S_b$ ($R_b = S_e^{-1}$). If R_1, S_1 are not multiples of a common word W then a finite sequence of such conjugations will yield $(R_2)_e \neq (S_2)_e$ and $(R_2)_e \neq (S_2)_b^{-1}$ without destroying (a) or (b). Property (a) is important in that if a pair (X, Y) satisfies (a) then it satisfies at most one of conditions $(T_1) - (T_5)$. Thus we need only find (R, S) satisfying some condition (T_K). Throughout the proof we do not mention the f_i's which are assumed to be chosen compatible with the pairs (R_i, S_i).

After Step (1) there are five distinct possibilities for (R_2, S_2): (i) $R_2 = 1$ or $S_2 = 1$ (ii) R_2 is not cyclically reduced (iii) $(R_2)_b = (S_2)_b$, $R_2 \neq 1$ (iv) $(R_2)_b = (S_2)_e^{-1}$, $R_2 \neq 1$ (v) none of (i)–(iv). Steps (2)-(5) are now applied in the following manner. If (i) occurs, apply Move (II) when $R_2 = 1$ so that $S_3 = 1$. Set $R_6 = R_3$ and $S_6 = S_3 = 1$ then (R_6, S_6) is of type T_1. If (ii) occurs, set $R_6 = R_2$, $S_6 = S_2$ then (R_6, S_6) is of type T_2. Likewise when (v) occurs. If (iii) occurs, either (R_2, S_2) is of type T_3 or an application of Move (II) yields $(R_3, S_3) = (\phi(S_2), \phi(R_2))$ of type T_3.

The remainder of the proof deals with the case when (iv) occurs. Here either (R_2, S_2) is of type T_4 or Move (II) yields (R_3, S_3) having type T_4 or $(R_3, S_3) = (L\overline{R}, L^{-1})$ for some L, $\overline{R} \in G$. It is this exceptional case to which Move (III) is applied. Let $(U, V) \equiv (R_4, S_4) =$ reduced form of $(\psi_\#(L), \psi_\#(\overline{R}))$. We analyze the various possibilities for (U, V) assuming (for each of writing) that in each case none of the previous cases applies.

Case 1. $(R_3)_e = (S_3)_b^{-1}$. From previous conditions, $(R_3, S_3) = (W^{K_1}, W^{K_2})$ where $K_1 > 0$ and $0 < -K_2 \le K_1$. Then $(U, V) =$ reduced form of $(Z^{|K_2|}, Z^{K_1 + K_2})$ where $Z = \psi_\#(W)$. Apply Step (4) to obtain (R_5, S_5) observing that $(R_5)_b = (S_5)_b$ and hence type T_3.

For the remaining cases we have (1.4.1): $L_b \ne \overline{R}_e^{-1}$, $L_e \ne \overline{R}_b^{-1}$, and $L_e \ne \overline{R}_e$.

Case 2. $U = 1$ or $V = 1$. In fact, must be true that $V = 1$. Clearly (R_6, S_6) is of type T_1.

Case 3. U is not cyclically reduced. Since L is cyclically reduced this only occurs when $L = a$---b or b^{-1}---a^{-1}. We only consider $L = a$---b; the other follows by taking inverses. From (1.4.1) we have that $\overline{R}_b \ne b^{-1}$ and $\overline{R}_e = a$ or b^{-1}. If $\overline{R} = a^{-1}$ or b then $V = p$---q where $p \ne a$, $q \ne a^{-1}$ and if $\overline{R}_b = a$ then $V = ab$---q, $q \ne a^{-1}$. Since $U = ab$---xa^{-1} where $x = b$ or a^{-1} it follows that (R_6, S_6) is of type T_2 or T_3.

Case 4. $U_b = V_e^{-1}$. From (1.4.1) this only occurs when $L_b \overline{R}_e = ab$ or $b^{-1}a^{-1}$ yielding $U = ab$--, $V = $--$xa^{-1}$ or $U = ax^{-1}$--, $V = $--$b^{-1}a^{-1}$ where $x = b$ or a^{-1}. If $V_b = a$ then apply Move (I) with $Q = a^{-1}$. Since U is cyclically reduced, $U_e \ne a^{-1}$ and we obtain (R_6, S_6) of type T_2 or T_3. If $V_b \ne a$ then $V_b \ne U_e^{-1}$ (as $L_e \ne \overline{R}_b^{-1}$) and we obtain type T_4 except when $x = 1$. But here $\overline{R} = b$ or $L = b^{-1}$ (depending on $L_b \overline{R}_e$) and by (1.4.1) we have $L_e = a$ or $\overline{R}_b = a^{\pm 1}$ respectively and so (R_4, S_4) has type T_5 except when $\overline{R}_b = a$ (type T_2).

Case 5. $U_e = V_b^{-1}$. This is the same as Case 4.

Case 6. V is not cyclically reduced. As Cases 1 through 5 do not occur simply apply Move (I) so that S_5 is cyclically reduced. Hence, type T_2.

Case 7. all others. This only leaves $U_b = V_b$, $U_e = V_e$, or neither. Apply Move (I) as necessary so that (R_5, S_5) has type T_2 or T_3. $\square$

The remainder of this paper is concerned with the proof of Theorem 1.1. In §2 we prove a Lemma which helps detect fixed points for certain maps. This is then applied throughout §§3, 4, and 5. Here, various Lemmas are given with the purpose of finding a good representative map $f \simeq F_{(X,Y)}$ achieving $MF(F_{(X,Y)})$ fixed points. The main result of these Lemmas is Proposition 5.8, which will enable us to describe the set $f^{-1}(A)$. In §6 all possibilities, up to isotopy, for $f^{-1}(A)$ are listed. Section

7 concludes the proof of the Theorem as the results of §§3–6 allow us to apply §2 and actually count $\#\text{Fix}(f)$.

2. Detection of fixed points. In this section we describe a method which will help "detect" fixed points, the main result of which is Lemma 2.1. This will be applied throughout the course of this paper. Serving as the key ingredient for a sequence of lemmas designed to find a good representative map, f, and then as the tool for counting $\#\text{Fix}(f)$. The reader is referred to (2.2) for a short discussion as to how this method and Lemma 2.1 are applied.

We begin with the setting for the Lemma. Let I denote the unit interval $[0,1]$ and let $p_i: I \times I \to I$, $i = 1$, 2, be the projection map $(x_1, x_2) \mapsto x_i$. Let $D \subset I \times I$ be a disk and let $f: \partial D \to I \times I$ be given so that $\text{Fix}(f) = \varnothing$ and the set $\mathscr{S} \equiv \mathscr{S}(f) = \{ x \in \partial D \mid p_1 \circ f(x) = p_1(x)\}$ is finite. We call $\mathscr{S}(f)$ the *detection set* for f.

Define a reduction procedure on $\mathscr{S}$ as follows: First, let $\mathscr{S}_0 = \{ x \in \mathscr{S} \mid$ there is a neighborhood $\mathcal{O}$ of x in ∂D such that $\mathcal{O} - x = U \cup V$ where $U \cap V = \varnothing$, $p_1 \circ f(u) > p_1(u)$ for $u \in U$, and $p_1 \circ f(v) < p_1(v)$ for $v \in V\}$. If T is a finite subset of ∂D then a *simple reduction* of T is the set $T - \{x, y\}$ such that (1) $\{x, y\}$ is the boundary of an arc $\alpha \subset \partial D$ with $\alpha \cap T = \partial \alpha$ and (2) $p_2 \circ f(x) > p_2(x)$ iff $p_2 \circ f(y) > p_2(y)$. A *reduction* of T is obtained by taking a sequence of simple reductions. The set T is *reduced* if no reductions exist.

LEMMA 2.1. *Let $f: \partial D \to I \times I$ be as above and let $\mathscr{S}'$ be a reduction of $\mathscr{S}_0$. If $\mathscr{S}'$ is reduced, then there exists a fixed point free extension of f to D iff $\mathscr{S}' = \varnothing$.*

REMARK. $\mathscr{S}'$ reduced implies that $\#\mathscr{S}'$ is an even integer. If $\tilde{f}: D \to I \times I$ is an extension of f then index $(\tilde{f}, \mathring{D}) = \pm \#\mathscr{S}'/2$. See [**J1**; p. 14] or [**B**; p. 59] for the fixed point index.

Proof. Using the reduction $\mathscr{S}'$ of $\mathscr{S}_0$ we define an extension $\tilde{f}$ of f which has exactly one fixed point. This is done in such a way that the index of this fixed point is easily calculated using Nielson Theory. The lemma is then established by invariance of the fixed point index.

To define $\tilde{f}$ first consider the special case when $\mathscr{S}' = \mathscr{S}_0 = \mathscr{S}$. Choose $p \in \mathring{D}$ and let $\{\gamma_w\}$, $w \in S^1$ (unit circle), be embeddings of I into D such that $\gamma_w(0) = p$, $\gamma_w(\mathring{I}) \cap \gamma_z(\mathring{I}) = \varnothing$ if $w \neq z$, $\gamma_w(\mathring{I}) \subset \mathring{D}$, $\gamma_w(1) \subset \partial D$, and $D = \bigcup_w \gamma_w(I)$. Set $\tilde{f}(p) = p$ and fix $w \in S^1$. Write $\gamma_w(t) = (U_t, V_t)$ and $f(\gamma_w(1)) = (U_1 + d, V_1 + e)$. Define $f'(\gamma_w(t)) = (U_t + t \cdot d, V_t + t \cdot e)$ and do for each $w \in S^1$. Set $\tilde{f} = r \circ f'$ where r is a retraction

of $R \times R$ onto $I \times I$. Clearly, $\tilde{f}$ extends f and since $\mathrm{Fix}(f) = \varnothing$ we have $d^2 + e^2 \neq 0$ and so $\mathrm{Fix}(\tilde{f}) = p$.

The general case can be done exactly as above but in order to illustrate the reduction procedure an alternate approach is given. First, pick $q \in \mathscr{S} - \mathscr{S}_0$ and let N be a regular neighborhood of q in D so that the arc $N \cap \partial D$ does not meet $\mathscr{S} - \{q\}$. Define $\tilde{f}$ on ∂N extending f so that $\mathscr{S}(\tilde{f} | \partial N) = \{q\}$ and then extend to N so that $p_1 \circ \tilde{f}(y) \neq p_1(y)$ for each $y \in \mathring{D}$. Hence, no fixed points. Do this for each point in $\mathscr{S} - \mathscr{S}_0$ to obtain $D' \subset D$ and $\tilde{f} \colon D - \mathring{D}' \to I \times I$ with $\mathscr{S}(\tilde{f} | \partial D') = \mathscr{S}_0$.

Now, let $\{x, y\}$ be the first pair of points removed in reducing $\mathscr{S}_0$ and α as in the definition. Let N be a regular neighborhood of α in D so that the arc $N \cap \partial D$ does not intersect $\mathscr{S}_0 - \{x, y\}$. Define $\tilde{f}$ on $\partial N - \alpha$ so that $\mathscr{S}(\tilde{f} | \partial N) = \{x, y\}$. Extend to N so that $\beta = \{y \mid p_1 \circ \tilde{f}(y) = p_1(y)\}$ is an arc and, by (2) in the definition, that $p_2 \circ \tilde{f}(y) \neq p_2(y)$ for each $y \in \beta$. Repeat for each simple reduction to obtain $\tilde{f} \colon D - \mathring{E} \to I \times I$ with $\mathscr{S}(\tilde{f} | \partial E) = \mathscr{S}'$. Complete the extension by applying the special case to E.

The advantage of defining $\tilde{f}$ in this manner is that in the special case it follows easily from the definition of the fixed point index that index $(\tilde{f}, p) = \pm \#\mathscr{S}'/2$. Notice that the points of $\mathscr{S}'$ alternate as to the direction of the second coordinate projection. As a result, a pair of points corresponds to one revolution about the unit circle.

To finish, if $\mathscr{S}' = \varnothing$ then index $(\tilde{f}, p) = 0$ and so $\tilde{f} \simeq \bar{f}$ rel ∂D with $\mathrm{Fix}(\bar{f}) = \varnothing$. If $\mathscr{S}' \neq \varnothing$ then by invariance of the fixed point index (rel ∂D), index $(g, \mathring{D}) \neq 0$ for any $g \simeq \tilde{f}$ rel ∂D and thus $\mathrm{Fix}(g) \neq \varnothing$. $\square$

REMARK 2.2. Application of reduction and Lemma 2.1. Consider the following general setting: Given $f \colon P \to P$ suppose that D is a disk contained in $C \equiv$ closure of a component of $P - A$ such that $f(D) \subset C$. In addition, suppose we can choose a product structure for C; a homeomorphism $h \colon C \to I \times I$, so that $\mathscr{S}(h \circ f \circ h^{-1} | h(\partial D)) \equiv \mathscr{S}$ is known and $\mathscr{S} = \mathscr{S}_0$ (also, finite). We set $\mathscr{S}(f) = h^{-1}(\mathscr{S})$ and to each point $p \in \mathscr{S}(f)$ we assign a *label*: '+' if $p_2 \circ h \circ f(p) > p_2 \circ h(p)$ and '−' otherwise. Then $\mathscr{S}(f)$ together with the labels acts as a detection set for $f | \partial D$. Reductions of $\mathscr{S}(f)$ using labels corresponding to reductions of $\mathscr{S}$.

3. Choosing a representative, I. This section deals with various Lemmas which are designed to improve the set $f^{-1}(A)$ for a given $f \colon P \to P$ (see Figure 1 for A). Lemmas 3.1–3.3 are essentially general position Lemmas while 3.4–3.7 are improvements in that they reduce $f^{-1}(A) \cap A$.

LEMMA 3.1. *Given $g \simeq F_{(X,Y)}$ there is a map f such that $\#\text{Fix}(f) \leq \#\text{Fix}(g)$ and $f \simeq F_{(X,Y)}$ rel ∂P.*

See Lemma 6.10 of [FH] or Lemma §3 of [J2] for a proof.

LEMMA 3.2. *Given $g \simeq F_{(X,Y)}$ rel ∂P there is a map $f \simeq g$ rel ∂P such that*
  (1) $\#\text{Fix}(f) \leq \#\text{Fix}(g)$.
  (2) $f^{-1}(A)$ *is a 1-dimensional proper submanifold of P transverse to A.*
  (3) $\text{Fix}(f) \cap f^{-1}(A) = \varnothing$.
  (4) *For each $p \in f^{-1}(A)$ and open set U_p containing p there exists an open set $\mathcal{O}_p$ with $p \in \mathcal{O}_p \subset U_p$ such that $f(\mathcal{O}_p)$ has nontrivial intersection with exactly two components of $P - A$.*

Proof. By general position there is a map h satisfying (1), (3), and $h^{-1}(A)$ is a 1-dimensional polyhedron whose vertices miss $A \cup \partial P$. Let Q denote the non-vertex points in $h^{-1}(A)$ which do not satisfy (4). Then Q consists of certain components of $h^{-1}(A)-\{\text{vertices}\}$. By a small homotopy the points of Q can be removed from $h^{-1}(A)$ without increasing the number of fixed points or altering h on $(h^{-1}(A) - Q) \cup \partial P$. Finish by altering h on a neighborhood of each vertex of $h^{-1}(A)$ as indicated in Figure 2 below. $\square$
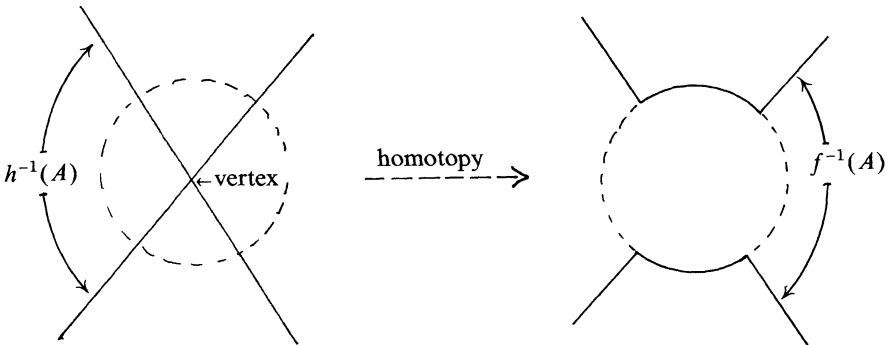

FIGURE 2

The content of the proof of Lemma 3.2 is that by a small adjustment of the map g we can replace the 0-dimensional non-manifold part of $g^{-1}(A)$ by the empty set. If the same process is applied to a homotopy between two maps then the 1-dimensional non-manifold set (given by general position) is replaced by a 0-dimensional set. Consequently,

LEMMA 3.3. *Suppose h_1, $h_2 \simeq F_{(X,Y)}$ rel ∂P with $h_i^{-1}(A)$ ($i = 0, 1$) being 1-dimensional submanifolds of P. Then there exist a homotopy H*

rel ∂P and a finite subset T of I such that $H_i = h_i$, $i = 0, 1$, and (i) if $t \notin T$ then $H_t^{-1}(A)$ is a 1-dimensional manifold and (ii) if $t \in T$ then $H_t^{-1}(A)$ has exactly one non-manifold point.

REMARK. Lemma 3.3 is only used in proving the Propositions in §6. Lemmas 3.1 and 3.2 are used throughout the paper and so we assume that, unless stated otherwise, all self-maps of P satisfy the conclusions of these two lemmas.

Let $f: P \to P$ be given. Choose embeddings, $\overline{A}_i$, $1 \le i \le 4$, of the unit interval $[0, 1]$ into P so that $\overline{A}_i[0, 1] = A_i$ and $\overline{A}_i(1) \subset \partial_3$ (see Figure 1). Given $x, y \in A_i$ we say that $x < y$ iff $\overline{A}_i^{-1}(x) < \overline{A}_i^{-1}(y)$ and $x > y$ iff $\overline{A}_i^{-1}(x) > \overline{A}_i^{-1}(y)$. We say x is inside (outside) of y when $x < y$ ($x > y$). Let $J_{x,y}$ denote the arc contained in A_i which has boundary $\{x, y\}$. Given $x < y$ and $f(\{x, y\}) \subset A_i$ we say that the pair of points $\{x, y\}$ is contracting if $f(x) > x$ and $f(y) < y$; expanding if $f(x) < x$ and $f(y) > y$.

A simple null curve is a subset, $\alpha(r, s)$, of $f^{-1}(A_r)$ homeomorphic to $[0, 1]$ with $\alpha(r, s) \cap A = \partial\alpha(r, s) \subset A_s$. A null curve is a connected curve formed by joining together a finite number of simple null curves; again written $\alpha(r, s)$. We write α in place of $\alpha(r, s)$ when either the values of r, s are clear or not of importance. Also, J_α is used in place of $J_{\partial\alpha}$. The null curve $\alpha(s, s)$ is contracting (expanding) if $\partial\alpha$ is contracting (expanding).

The following notation will be used throughout the remainder of this section and into the next. Let $\alpha(r, s)$ be a simple null curve and let P_L, P_C, P_R be as in Figure 1. Let D_α denote the disk with boundary $\alpha \cup J_\alpha$, $R_\alpha = P_z$ where $z \in \{L, C, R\}$ so that $D_\alpha \subset R_\alpha$, and $L_\alpha = P_w$ where $w \ne z$ and $J_\alpha \subset L_\alpha$.

LEMMA 3.4. (Simple Crossing Reduction) Suppose $\{x, y\} \subset A_s \cap f^{-1}(A_r)$ with $\mathring{J}_{x,y} \cap f^{-1}(A) = \varnothing$. Then there is a map $g \simeq f$ rel ∂P such that $\mathrm{Fix}(g) \subset \mathrm{Fix}(f)$ and (i) if $r = s$ then $g^{-1}(A) \cap A = f^{-1}(A) \cap A$ with $\{x, y\}$ being the boundary for a simple null curve in $g^{-1}(A)$, (ii) if $r \ne s$ then $g^{-1}(A) \cap A = (f^{-1}(A) \cap A) - \{x, y\}$.

Proof. If $r = s$ let λ be an arc parallel to $J_{x,y}$ contained in the component of $P - A$ which does not meet $f(\mathring{J}_{x,y})$, $\partial\lambda \subset f^{-1}(A)$. If $r \ne s$ let $\lambda = J_{x,y}$ (see Figure 3 below). Let (N, N_0) be a regular neighborhood of $(\lambda, \partial\lambda)$ in $(P, f^{-1}(A))$ such that if $r = s$ then $N \cap A = \varnothing$ and if $r \ne s$ then $N \cap A_s$ is an arc. Also, $f(N) \cap N = \varnothing$. Choose arcs β_1, β_2 in N such that $\partial\beta_i \subset N_0$ and $\mathring{\beta} \cap (f^{-1}(A) \cup \lambda \cup \partial N) = \varnothing$. Define g so that $g = f$ on $P - \mathring{N}$ and $g(\beta_i) \subset A_r$. Now, if N is chosen close enough to λ

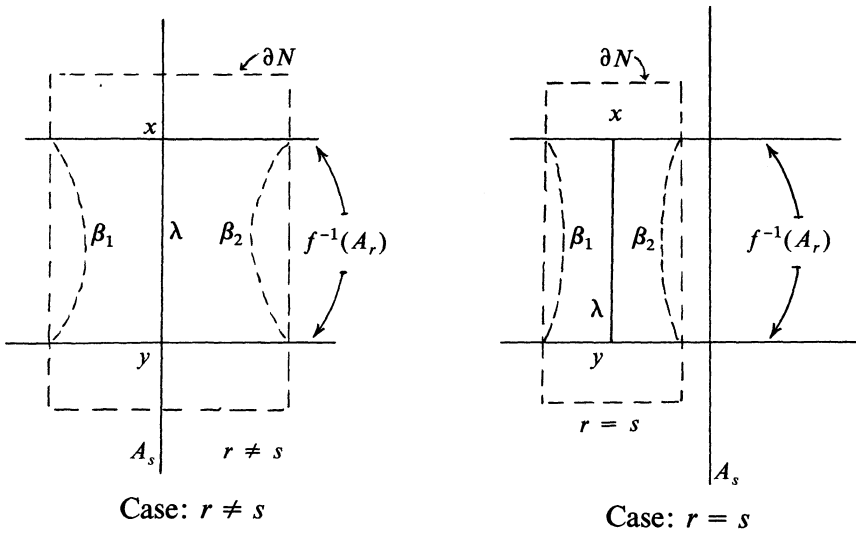Case: $r \ne s$ Case: $r = s$

FIGURE 3. λ in Lemma 3.4

then we extend to each component of $N - (\beta_1 \cup \beta_2)$ by mapping into $P - (N \cup A)$ (this is simply due to the choice of λ). As a result $g(N) \cap N = \varnothing$ and $g^{-1}(A) = (f^{-1}(A) - N) \cup \beta_1 \cup \beta_2$.   □

LEMMA 3.5. (*Null Curve Reduction.*) *Given* $f: P \to P$ *and* $\alpha(r, s)$ *a simple null curve in* $f^{-1}(A)$, *there is a map* $g \simeq f$ *rel* ∂P *such that* $\#\mathrm{Fix}(g) \le \#\mathrm{Fix}(f)$ *and*

$$g^{-1}(A) \cap A = \begin{cases} \left(f^{-1}(A) \cap A\right) - \mathring{J}_\alpha & \text{if } r = s, f(J_\alpha) \subset L_\alpha, \text{ and } \alpha \text{ is} \\ & \text{expanding or contracting.} \\ \left(f^{-1}(A) \cap A\right) - J_\alpha & \text{otherwise.} \end{cases}$$

Proof. First consider the special case when $\mathring{J}_\alpha \cap f^{-1}(A) = \varnothing$. Here we need to remove $\partial\alpha$ from $f^{-1}(A) \cap A$ if either $r \ne s$, $f(J_\alpha) \subset R_\alpha$, or α is neither expanding nor contracting. If $r \ne s$, just use the proof of Lemma 3.4 with $\lambda = J_\alpha$. If $r = s$ and α is neither expanding nor contracting then $f \simeq f_1$ with support on a neighborhood of $\partial\alpha$ so that $f_1(\partial\alpha) = \{p\}$ where $p \notin J_\alpha$. Mimic the proof of 3.4 with $\lambda = J_\alpha$. Noting that since $f_1(\lambda) \cap \lambda = \varnothing$ we can choose N so that $N \cap f_1(N) = \varnothing$ and homotope so that $g(N) \cap N = \varnothing$. Hence $\#\mathrm{Fix}(g) \le \#\mathrm{Fix}(f)$ as desired. To finish this special case we now assume that $r = s$, $f(J_\alpha) \subset R_\alpha$, and α is expanding or contracting.

Let N be a regular neighborhood of D_α so that $N \cap A_s$ and $\bar{\alpha} = N \cap f^{-1}(A)$ are proper arcs in N. Choose an arc β in N with $\beta \cap A_s = \varnothing$ and $\partial\beta = \partial\bar{\alpha}$. Write $\partial N = \gamma_1 \cup \gamma_2$ with $\gamma_1 \cap \gamma_2 = \partial\bar{\alpha}$ and $\gamma_1 \subset L_\alpha$. Define g so that $g = f$ outside of N, $g(\beta) \subset A_s$, g maps the disk bounded

by $\beta \cup \gamma_1$ into R_α (by hypothesis $f(\gamma_1) \subset R_\alpha$), and the remaining disk is mapped into $\mathring{L}_\alpha$ having at most one fixed point. To achieve the last condition map $N \cap A_s$ into $\mathring{L}_\alpha$ and extend to the two disk components. On the component between β and $N \cap A_s$ there must be a fixed point so we choose an extension which has exactly one fixed point and maps the interior into $\mathring{L}_\alpha$. The remaining disk is simply mapped into $\mathring{L}_\alpha$ as its boundary is.

Clearly, $g \simeq f$ and $g^{-1}(A) = (f^{-1}(A) - \bar{\alpha}) \cup \beta$. To see that $\# \text{Fix}(g)$ $\leq \# \text{Fix}(f)$ we need to show that $f|D_\alpha$ has a fixed point. To do so choose a product structure for R_α so that A_s corresponds to $0 \times I$ (see 2.2). It is easy to see that $\mathscr{S}(f) = \partial\alpha$ and since $\partial\alpha$ is either expanding or contracting these two points have opposite labels. Hence, $\mathscr{S}$ is reduced and, by Lemma 2.1, we have a fixed point.

A second special case to consider is when

$$J_\alpha \cap f^{-1}(A) = \{ y_0 < x_1 < y_1 < x_2 < \cdots < y_{n-1} < x_n < y_n < x_{n+1} \}$$

where $\{ x_i, y_i \} = \partial\delta_i(s, s)$ with $D_{\delta_i} \subset D_\alpha$ and $f(D_{\delta_i}) \subset L_\alpha$. We also assume that either each δ_i is contracting or each δ_i is expanding and that if $r = s$ the pairs $\{ y_0, x_1 \}$ and $\{ y_n, x_{n+1} \}$ are either expanding or contracting.

First apply Lemma 3.4 to each pair $\{ y_i, x_{i+1} \}$ ($1 \leq i < n$ if $r \neq s$ and $0 \leq i \leq n$ if $r = s$) so that the pair of points bounds a simple null curve, δ_i', contained in L_α. Now, if $r = s$ let N be a regular neighborhood of $D_\alpha \cup \bigcup_{i=0}^n D_{\delta_i'}$ chosen so that $N \cap A_s$ is an arc and $\partial N \cap f^{-1}(A) = \varnothing$ (Figure 4). Define g so that $g = f$ outside of N and since $f(D_{\delta_i}) \subset L_\alpha$ extend to N by mapping into $\mathring{L}_\alpha$ so that $N \cap L_\alpha$ contains exactly one fixed point.

If $r \neq s$ choose N so that $\text{Fix}(f|\partial N) = \varnothing$, $N \cap A_s$ is an arc, and $N \cap f^{-1}(A) = \bar{\alpha} \cup \bar{\delta}$, arcs containing α and $\cup\delta_i$ respectively (Figure 4). Choose proper arcs β_1, β_2 in $N \cap L_\alpha$ so that $\partial\beta_1 = \partial\bar{\delta}$ and $\partial\beta_2 = \partial\bar{\alpha}$. Define g so that $g = f$ outside of N, $g(\beta_1) \subset A_s$, $g(\beta_2) \subset A_r$, and $g(N \cap A_s) \subset L_\alpha - N$. Extend to the domain between β_1 and β_2 by mapping into R_α and to $N \cap R_\alpha$ by mapping into L_α. As $g(N \cap A_s) \cap N$ $= \varnothing$ we can extend to the domain between β_2 and $N \cap A_s$ missing N (assuming ∂N is "near" $\bar{\alpha}$). Finally, the remaining domain is mapped into L_α so as to have at most one fixed point.

This case is completed by showing that $f|D_\alpha - \cup D_{\delta_i}$ has a fixed point. Choose a product structure for R_α so that A_s corresponds to $0 \times I$ and if $r \neq s$ then A_r corresponds to $1 \times I$. Then $\mathscr{S}(f) = \bigcup_{i=1}^n \partial\delta_i$ ($r \neq s$) or $\bigcup_{i=0}^n \partial\delta_i'$ ($r = s$). By hypothesis the labels alternate and so $\mathscr{S}$ is reduced and nonempty.
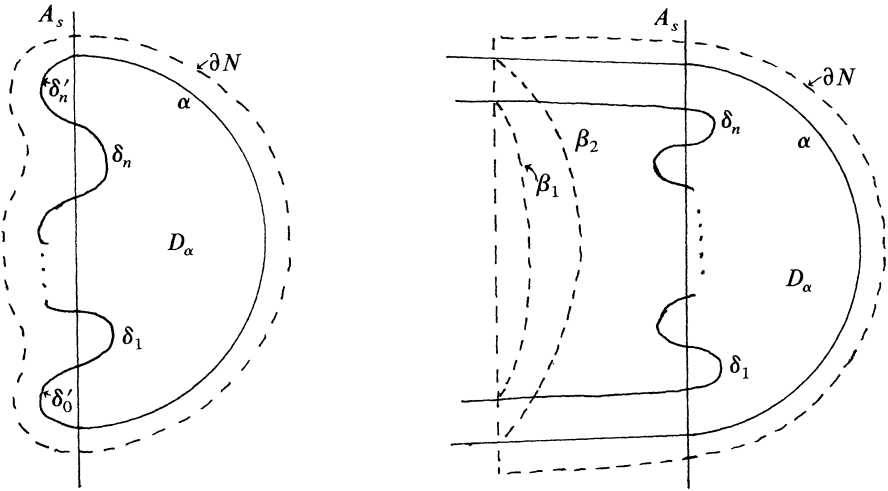
FIGURE 4

REMARK. In the case $r \neq s$ it is quite possible that $\text{Fix}(g \mid N) = \varnothing$. This depends on $f \mid (\partial N \cap L_\alpha)$ which determines whether or not $f \mid N \cap L_\alpha$ has a fixed point which cancels with the one in D_α.

The general case is done by first applying Lemma 3.4 to pairs of points in J_α. Next use the first special case where possible so that now if $\mathring{J}_\alpha \cap f^{-1}(A) \neq \varnothing$ then there exists α' and $\delta_1, \ldots, \delta_n$ as in the second special case. Remove these null curves and repeat the whole process continuing until all curves (except possibly α) are removed from J_α. $\square$

LEMMA 3.6. *Suppose C is a simple closed curve contained in $f^{-1}(A)$ with $C \cap A = \varnothing$. Then there is a map $g \simeq f$ rel ∂P such that $\text{Fix}(g) \subset \text{Fix}(f)$, $g^{-1}(A) \subset f^{-1}(A)$, and $C \cap g^{-1}(A) = \varnothing$.*

Proof. Since $C \cap A = \varnothing$ there is a disk $D \subset P - A$ with $\partial D = C$. Let N be a regular neighborhood of D which misses A and satisfies $f(\partial N) \cap N = \varnothing$ (because $f(C) \subset A$). Let $E \subset P$ be a disk which contains $f(\partial N)$ and does not intersect $N \cup A$. Set $g = f$ on $P - \mathring{N}$ and extend by mapping N into E. Thus, $\text{Fix}(g) = \text{Fix}(f) \cap (P - N)$ and $g^{-1}(A) = f^{-1}(A) \cap (P - N)$ as desired. $\square$

COROLLARY 3.7. *If $f: P \to P$ satisfies the conclusions of Lemmas 3.4, 3.5, 3.6 then $f^{-1}(A)$ does not contain any inessential simple closed curves.*

Proof. Suppose $C \subset f^{-1}(A_r)$ bounds a disk D. As $D \cap A$ is a finite collection of proper arcs in D it follows from Lemma 3.4 that $C \cap A_s = \varnothing$ when $s \neq r$. Thus $C = \alpha_1 \cup \cdots \cup \alpha_n$ where $\alpha_i(r, r)$ is a simple null curve.

By applying Lemma 3.5 we reduce to $n = 2$ and noting that $f(J_{\alpha_i}) \subset R_{\alpha_i}$ for one such i we have that $C \cap A = \varnothing$. But this contradicts Lemma 3.6. □

4. Analysis of curves. One of the purposes of the simple crossing reduction and null curve reduction lemmas is to gain a better understanding of maps $f: P \to P$ by concentrating on the geometric intersection of $f^{-1}(A)$ with A. Notice that both of these lemmas tend to reduce this intersection. Continuing along these lines we wish to explore other ways in which the geometric intersection of $f^{-1}(A)$ with A may be reduced. It is with this in mind that the following notation for certain curves in P is introduced.

Let γ be a curve in P transverse to A such that $\partial\gamma \subset \partial P \cup A$, $\gamma \cap A \neq \varnothing$, and $\mathring{\gamma} \cap \partial P = \varnothing$. Choose $h: [0,1] \Rightarrow \gamma$ so that $h|_{(0,1)}$ is one-to-one and $\gamma \cap (A \cup \partial P) = h(\bigcup_{k=0}^{n} k/n)$ for some $n \geq 1$. We use the notation $(x_0, x_1, \ldots, x_n)$ to describe the curve γ where $x_k \in \{1, 2, 3, 4, \Delta_1, \Delta_2, \Delta_3\}$ and $x_k = r$, $r \in \{1, 2, 3, 4\}$ iff $h(k/n) \subset A_r$, $x_k = \Delta_i$, $i \in \{1, 2, 3\}$ iff $h(k/n) \subset \partial_i$ (see Figure 1 for ∂_i). If $\gamma_1 = (x_0, \ldots, x_n)$, $\gamma_2 = (y_0, \ldots, y_m)$ with $\partial\gamma_1 \cap \partial\gamma_2 \neq \varnothing$ (say $h_1(1) = h_2(0)$) then we may denote their union by any of $(x_0, \ldots, x_n, y_1, \ldots, y_m)$, $(x_0, \ldots, x_{n-1}, \gamma_2)$, $(\gamma_1, y_1, \ldots, y_m)$, or (γ_1, γ_2).

REMARKS. (1) As orientations are not considered γ can be expressed using either $(x_1, \ldots, x_n)$ or $(x_n, \ldots, x_1)$. For example, $(1, 2)$ and $(2, 1)$ may represent the same curve.

(2) Given $(x_1, \ldots, x_n)$ by no means do we determine a unique curve (up to isotopy, say) in P (see Figure 5). For most of our purposes this will not pose a problem but whenever necessary additional information will be given so as to avoid confusion.

If C is a disjoint collection of curves in P we assign to each curve a *μ-value*: A function $\mu: C \to \{1, 2, 3, 4\}$ such that if $\bigcup C = f^{-1}(A)$ for some f then $\mu(\gamma) = r$ iff $\gamma \subset f^{-1}(A_r)$ for each $\gamma \in C$. Throughout this paper, if the curve γ under consideration is clear we may write $\mu = r$ in place of $\mu(\gamma) = r$. For the following definitions we assume that the curve $\gamma = (x_1, \ldots, x_n)$ is a subset of $f^{-1}(A)$ for some $f: P \to P$. If $\gamma \subset \mathring{P}$ then there is a unique curve $\bar{\gamma} = (x_0, x_1, \ldots, x_n, x_{n+1})$ such that the closure of $\bar{\gamma} - \gamma = (x_0, x_1) \cup (x_n, x_{n+1})$. The curve $\bar{\gamma}$ is called the *simple extension* of γ in $f^{-1}(A)$. The simple extension of γ can also be expressed in the form (x_0, γ, x_{n+1}). In general, an *extension* of γ is any curve of the form $(\delta_1, \gamma, \delta_2)$. In many instances we will work with "one-sided" extensions, i.e. when one of δ_1 or δ_2 is degenerate.
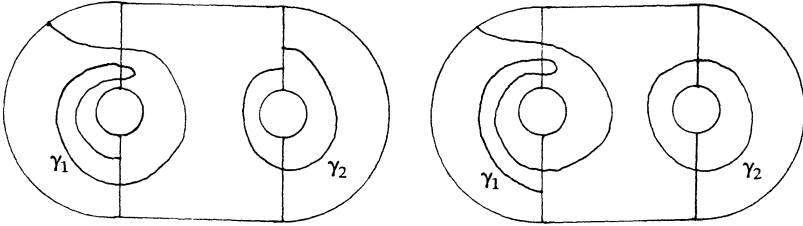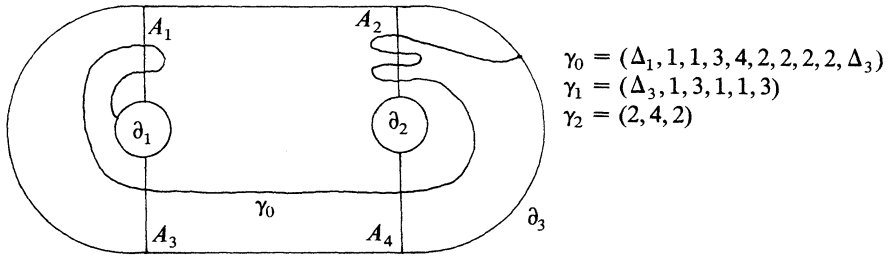
$$\gamma_0 = (\Delta_1, 1, 1, 3, 4, 2, 2, 2, 2, \Delta_3)$$
$$\gamma_1 = (\Delta_3, 1, 3, 1, 1, 3)$$
$$\gamma_2 = (2, 4, 2)$$

FIGURE 5. Examples

We say that $\gamma = (x_1, \ldots, x_n)$ is a *turn* if $x_i = x_j$ for all i, j, n is even, and the simple extension (x_0, γ, x_{n+1}) of γ satisfies $x_0 \neq x_1$ and $x_n \neq x_{n+1}$. The curve γ is a *crossing* if in the above definition n is odd and in addition, $\mu(\gamma) = x_1$. A turn or crossing is *simple* iff $n \leq 2$. Notice that turns and crossings are finite unions of null curves. Simple turns are in fact null curves and simple crossings are degenerate curves. As with null curves we let $J_\gamma \equiv J_{\partial \gamma} \equiv$ arc of $A_{\mu(\gamma)}$ joining the points of $\partial \gamma$.

Suppose that $f^{-1}(A)$ satisfies the conclusions of Lemmas 3.4 and 3.5. If γ is a turn in $f^{-1}(A)$ then it follows that $\gamma \cap A_r \neq \varnothing$ iff $r = \mu(\gamma)$. If γ is either a turn or a crossing then $J_\gamma \cap \gamma = \{ p_1 < p_2 < \cdots < p_n \}$ where each pair of points $\{ p_i, p_{i+1} \}$ is either contracting or expanding. If γ is a turn we say γ is *contracting* (*expanding*) if $\{ p_1, p_2 \}$ is contracting (expanding). If γ is a crossing then it is *contracting* if $f(p_1) < p_1$ and *expanding* if $f(p_1) > p_1$.

A *critical region*, R, for f is the closure of a component of $P - (A \cup f^{-1}(A))$ such that $f(\mathring{R})$ is contained in the component of $P - A$ which contains $\mathring{R}$. It is a result of Corollary 3.7 that all critical regions are disks. Given a curve γ we say that γ *meets* R (vice versa) if $\gamma \cap R \neq \varnothing$.

Finally, the phrase "f can be adjusted so that/by..." means that there is a map $g \simeq f$ rel ∂P with $\# \text{Fix}(g) \leq \# \text{Fix}(f)$ and g satisfies the conditions indicated in the phrase.

LEMMA 4.1. *Let R be a critical region which meets only the curves γ_1, $\gamma_2 \subset f^{-1}(A)$. Let $x = \mu(\gamma_1)$ and let δ, δ' denote crossings. Let C denote the component of $P - A$ containing R. If*

(4.1.1) $\gamma_1 = (\Delta_i, \delta)$, $\gamma_2 = (\Delta_i, x)$;   *or*

(4.1.2) $\gamma_1 = (\delta, y)$, $\gamma_2 = (x, y)$ *with $1 \le y \le 4$ and $\mu(\gamma_2) \ne y$; or*

(4.1.3) $\gamma_1 = (\delta, y)$, $\gamma_2 = (x, \delta')$ *with $\mu(\gamma_2) = y$*

then f can be adjusted so that in (4.1.1) δ is simple and contracting when $i = 3$, and simple and expanding otherwise; in (4.1.2) δ is simple and $f(\delta \cap A_x)$ is contained in the component of $A_x - \delta$ which intersects the component of $C - \delta$ containing $A_{\mu(\gamma_2)}$; in (4.1.3) $\#(\delta \cap A_x) = \#(\delta' \cap A_y)$ and $f(\partial \delta)$ is in the same component of $C - R$ as $f(\partial \delta')$.

Proof. In (4.1.1) choose a product structure for C such that A_x, $A_{\mu(\gamma_2)}$, $\Delta_i \cap C$ correspond to $0 \times I, 1 \times I, I \times 1$ respectively and $\mathscr{S}(f) = (\delta \cap A_x) \cup \{p\}$ where $p \in \Delta_i \cap R$. By Lemma 3.5, the points in $\delta \cap A_x$ alternate expanding and contracting and so the points of $\partial \delta$ are all labeled the same. If this label agrees with that of p then $\#\mathscr{S}' = \#(\delta \cap A_x) - 1$, otherwise we get $\#(\delta \cap A_x) + 1$. By Lemma 2.1, unless the labels of $\partial \delta$ and p agree and δ is simple we have a fixed point in R. As each crossing meets exactly two critical regions we do not increase the total number of fixed points if we adjust f to make δ simple and the labels match. It now follows that $\delta \cap A_x$ is contracting when $i = 3$ and expanding otherwise.

In (4.1.2) choose a product structure for C such that A_x, A_y, $A_{\mu(\gamma_2)}$ correspond to $0 \times I$, $1 \times I$, subset of $I \times 0$ respectively and $\mathscr{S}(f) = (\delta \cap A_x) \cup \{p\}$ where $p \in \gamma_2$. As in (4.1.1) we can make δ simple with $\delta \cap A_x$ having the same label as p. This establishes (4.1.2).

In (4.1.3) we arrange that $\mathscr{S} = (\delta \cap A_x) \cup (\delta' \cap A_y)$ and so no fixed point occurs in R under the conclusion of (4.1.3). Notice that we can not make both δ and δ' simple as this may increase the number of fixed points.   □

LEMMA 4.2. *Suppose $f^{-1}(A)$ contains the curve $\bar{\alpha} = (\Delta_3, \alpha, x)$ where α is a turn. Let R be the critical region meeting $\bar{\alpha}$ and $\sigma = (\Delta_3, \delta)$ where δ is a crossing at A_x. If δ is simple—contracting then we can adjust f by pulling the turn α off of $A_{\mu(\alpha)}$.*

Proof. Set up a product structure for the component of $P - A$ containing R so that $A_{\mu(\alpha)}$, A_x correspond to $0 \times I$, $1 \times I$ respectively
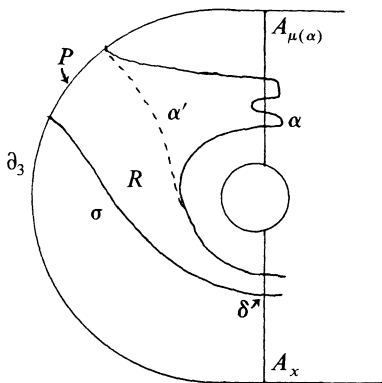
FIGURE 6. Proof of 4.2

and $\mathscr{S}(f) = (\alpha \cap A_{\mu(\alpha)}) \cup (\delta \cap A_x) \cup \{p\}$, $p \in \partial_3$. Since δ is simple-contracting both p and $\delta \cap A_x$ are labeled the same. Thus, $\mathscr{S}' = \alpha \cap A_{\mu(\alpha)}$ is reduced and so R contains a fixed point. By a homotopy with support on a neighborhood of α we replace $\bar{\alpha}$ by $\alpha' = (\Delta_3, x)$. As the resulting critical region (corresponding to R) need not contain a fixed point the total number of fixed points has not increased.   □

LEMMA 4.3. *Suppose* $f^{-1}(A)$ *contains the curve* $\bar{\alpha} = (\Delta_i, \alpha, x)$, $i \in \{1, 2\}$, *where* α *is a turn. Let* R *be the critical region meeting* $\bar{\alpha}$ *and* $\sigma = (\Delta_i, \delta)$ *where* δ *is a crossing at* A_x. *If* δ *is simple-expanding then we can adjust* f *by pulling the turn* α *off of* $A_{\mu(\alpha)}$.

Proof. Same as Lemma 4.2.   □

LEMMA 4.4. *Let* α *be a turn in* $f^{-1}(A)$ *with* $\mu(\alpha) = x$. *Let* γ_1, γ_2 *be* (x, y) *curves with* $\mu(\gamma_1) \neq \mu(\gamma_2)$ *and* (x, y) *being one of* $(1, 3)$, $(2, 4)$, $(1, 2)$, *or* (3.4). *Suppose* f *has a critical region,* R, *meeting* α, γ_1, *and* γ_2 *but no other curves. If* $\mu(\gamma_j) = y$ *implies that the crossing at* A_y *is simple and maps into the component of* $P_C - R$ *which contains* $A_{\mu(\gamma_{j'})}$, $j' \neq j$, *then we can adjust* f *by pulling the turn* α *off of* A_x.

Proof. Because $\mu(\gamma_1) \neq \mu(\gamma_2)$ the region R is contained in P_C. Set up a product structure for P_C so that A_x, A_y correspond to $0 \times I$, $1 \times I$ respectively and $\mathscr{S}(f) = (\alpha \cap A_x) \cup \{p_1, p_2\}$ where $p_j \subset \gamma_j$. If $\mu(\gamma_j) = y$, then p_j is the crossing at A_y and by hypothesis is labeled the same as $p_{j'}$. If $\mu(\gamma_j) \neq y$ for each j, the condition that $(x, y) \neq (1, 4)$ or $(2, 3)$ implies that p_1, p_2 have the same label. Hence, $\mathscr{S}' = \alpha \cap A_x$ so we may remove α without increasing the number of fixed points.   □

Given a collection of curves C let C' denote the collection of curves obtained from C by replacing each turn and crossing in C by a simple turn or crossing. Thus, the turn $(x_1, \ldots, x_n)$ is replaced by (x_1, x_n) and the simple extension $(x, y_1, \ldots, y_m, z)$ of a crossing is replaced by (x, y_1, z). This replacement does not, in general, help reduce the number of fixed points for a map f $(C = f^{-1}(A))$ but will be useful when dealing with extensions of curves.

Let $\Omega(C) = \{\alpha \mid \alpha$ is a turn in $C\}$. There is a natural correspondence between $\Omega(C)$ and $\Omega(C')$ so that given $\alpha \in \Omega(C)$ we let α' denote the corresponding element of $\Omega(C')$.

Define a relation $<$ on $\Omega(C)$ as follows: Given α, $\beta \in \Omega(C)$ then $\alpha < \beta$ iff there exists a sequence $\beta_0', \beta_1', \ldots, \beta_n'$ of curves satisfying: (1) $\beta' = \beta_0'$ (2) for each i, β_i' is a simple extension of β_{i-1}' in C' such that $\partial \beta_i' \subset A_r$ for some r and β_i' contains exactly one turn, namely β_0' (3) α' is contained in the disk bounded by $\beta_n' \cup J_{\partial \beta_n'}$.

We use the notation $\alpha \ll \beta$ if the above definition holds with $n = 1$. It follows from Lemma 3.4 that if $\beta \in \Omega(f^{-1}(A))$ has a simple extension of the form (x, β, x), $1 \le x \le 4$, then there is a unique turn $\alpha \in \Omega(f^{-1}(A))$ with $\mu(\alpha) = x$ and $\alpha \ll \beta$.

See Figure 7 for examples illustrating $<$ and $\ll$. The example $\delta \ll \gamma \ll \delta$ displays one of the problems which may arise—"cyclic behavior". In order to avoid problems caused by this behavior we have the following lemma.

LEMMA 4.5. (*Spiral Elimination.*) *Suppose* $f: P \to P$ *satisfies the conclusions of all previous lemmas. Then* f *can be adjusted so that given any sequence* $\alpha_1, \ldots, \alpha_m$ *in* $\Omega(f^{-1}(A))$ *with* $\alpha_i < \alpha_{i+1}$ *for each* i, *then* $\alpha_1 \ne \alpha_m$.

FIGURE 7. $\alpha < \beta$ and $\delta \ll \gamma \ll \delta$
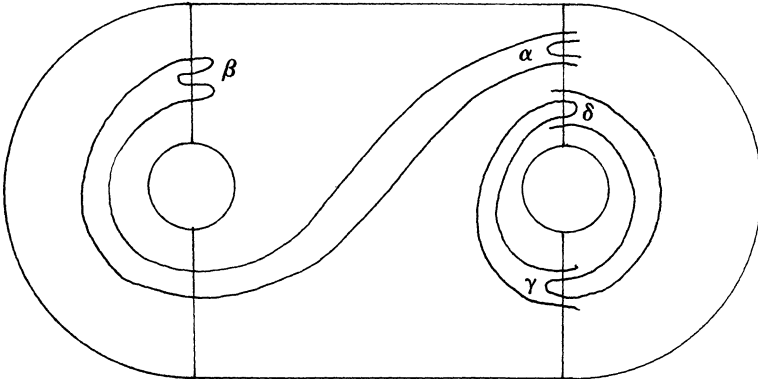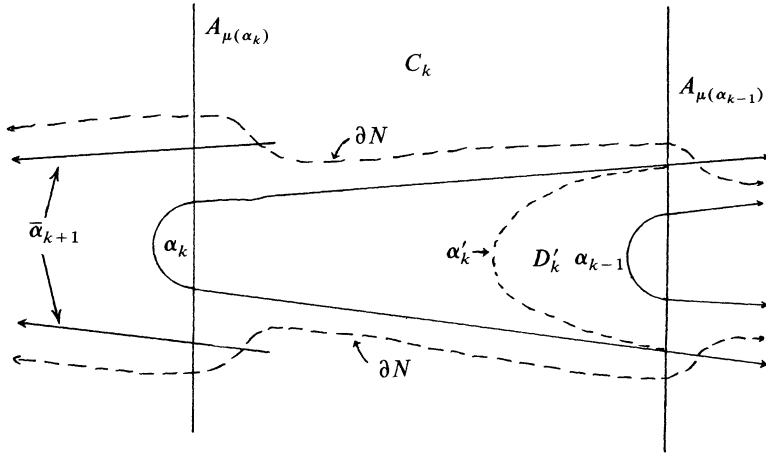
FIGURE 8

Proof. Consider the special case $\alpha_1 \ll \alpha_2 \ll \cdots \ll \alpha_n \ll \alpha_1$. For each k, let $\bar{\alpha}_k$ be the simple extension of α_k; α'_k the curve obtained by pulling $\bar{\alpha}_k$ off of $A_{\mu(\alpha_k)}$; D_k the disk bounded by $\bar{\alpha}_k \cup J_{\bar{\alpha}_k}$; D'_k the disk bounded by $\alpha'_k \cup J_{\alpha'_k}$; N a regular neighborhood of $\cup D_k$ chosen close to $\cup D_k$; C_k the component of $P - A$ containing D'_k; and E_k the component of $N - A$ which contains D'_k. See Figure 8 for an illustration.

Define a map g by setting $g = f$ on $P - \mathring{N}$ and so that $g(\alpha'_k) \subset A_{\mu(\alpha_k)}$. By Lemma 3.5 it follows that $f(D_k - D_{k-1}) \subset C_k$ and as $D_{k'}$ is to replace $D_k - D_{k-1}$ we extend g to D'_k by mapping into $(C_k - D'_k)$. Finally, extend to $E_k - D'_k$ by mapping into the appropriate component of $(P - A) - C_k$. Note that $\text{Fix}(g) \cap N = \varnothing$. Now reduce $g^{-1}(A)$ by applying Lemmas 3.4 through 4.4 obtaining h. If $h^{-1}(A)$ contains $\beta_1 \ll \cdots \ll \beta_k \ll \beta_1$ repeat the above procedure noting that since $g^{-1}(A) \cap A = (f^{-1}(A) \cap A) - \cup \alpha_k$ then $h^{-1}(A) \cap A$ is a proper subset of $f^{-1}(A) \cap A$ and so we need only repeat a finite number of times. To see that the result is the desired map we finish by showing that $\alpha_1 < \cdots < \alpha_1$ implies that there are turns $\alpha_2, \ldots, \alpha_n$ so that $\alpha_1 \ll \alpha_2 \ll \cdots \ll \alpha_n \ll \alpha_1$.

The cyclic behavior of $<$ is closely related to isotopy classes of simple closed curves in P. This relationship is obtained by collapsing a turn γ to the midpoint of J_γ and when $\alpha < \beta$ the extensions of β which define $\alpha < \beta$ collapse to an arc which joins the midpoints of J_α and J_β. As there are only three simple closed curves in P; namely $(2,4,2)$, $(1,3,1)$, and $(1,2,4,3,1)$ the possibilities for curves $\alpha_1 < \cdots < \alpha_1$ are limited. Now, starting with $\alpha = \alpha_1$ let λ_1, λ_2 be the two curves in R_α (see §3) which meet $A_{\mu(\alpha)}$ adjacent to α. Since $\alpha_1 < \cdots < \alpha_1$ these curves exist and $\lambda_i = (\mu(\alpha), y)$ for some y. From the above discussion each of λ_1, λ_2

collapses to an arc contained in a simple closed curve and thus $(\mu(\alpha), y)$ is one of $(1, 3), (2, 4), (1, 2),$ or $(3, 4)$. By Lemma 4.4 (if necessary) we have $\mu(\lambda_i) = y$ and by Lemma 3.4, λ_1 and λ_2 must join to form α' with $\alpha \ll \alpha'$. Set $\alpha_2 = \alpha'$ and repeat this process using $\alpha = \alpha_2, \alpha_3, \ldots$ until the resulting α' is equal to α_1. $\square$

As a consequence of Lemma 4.5 we can partition $\Omega(f^{-1}(A))$ into chains $\alpha_1 < \alpha_2 < \cdots < \alpha_m$ where $\alpha_i = \alpha_j$ iff $i = j$ or into maximal chains $\beta_1 \ll \beta_2 \ll \cdots \ll \beta_n$ where $\beta_i = \beta_j$ iff $i = j$. In the latter case β_1 is a *minimal* turn and β_n is a *maximal* turn. It follows from Lemma 3.4 that the boundary of the simple extension of β_1 is not contained in some A_r. Notice that as β_n is maximal there is no turn γ satisfying $\beta_n \ll \gamma$ but it is quite possible that there is a turn α with $\beta_n < \alpha$. In the next section we will see that this need not occur (although it is quite feasible). In fact, we can arrange that Ω has at most one nontrivial chain—consisting of at most four turns.

5. Choosing a representative, II. The main objective of this section is Proposition 5.8 which, when used together with the results of Section 6, gives a description of $C_f \equiv (f^{-1}(A))'$ (see §4). This is done by; first, Lemmas 5.1, 5.2 which relate C_f and $f^{-1}(A)$, and then Lemmas 5.3–5.7 which help reduce $\Omega(f^{-1}(A))$. Besides being used in the proof of 5.8 these lemmas will appear again in §7. For each lemma, we assume that $f: P \to P$ is given satisfying the conclusions of all previous lemmas.

A *full Δ_i-spiral* is a curve, γ, containing no full turns, with $\partial\gamma$ either empty or contained in some A_r and $\gamma \cup J_{\partial\gamma}$ is a simple closed curve in P isotopic to ∂_i. A *Δ_i-spiral* is either a full Δ_i-spiral or a curve $\gamma = \gamma_0 \cup \gamma_1$ where γ_0 is either empty or a proper subcurve of a full Δ_i-spiral and $\gamma_1 = (x, \Delta_i)$. In the latter case $J_{\partial\gamma}$ denotes the arc of A_r joining $\partial\gamma_0 - \partial\gamma_1$ to ∂_i. Two Δ_i-spirals γ and γ' are *adjacent* if $\partial\gamma \cap \mathring{P}$ and $\partial\gamma' \cap \mathring{P}$ are contained in the same A_r and no curves occur in the region between the two. Let $A(\gamma)$ denote the union of all arcs, I, in A satisfying $\partial I \subset \partial_i \cup \gamma$ and $\mathring{I} \cap \gamma = \varnothing$. Let $C(\gamma)$ denote the component of $P - (\gamma \cup J_{\partial\gamma})$ which does not intersect ∂_j for all $j \neq i$.

LEMMA 5.1. *Let γ_j, $1 \leq j \leq 4$, be disjoint Δ_3-spirals such that γ_j is adjacent to γ_{j+1} with γ_1 nearest ∂_3 and $\partial\gamma_j \cap \mathring{P} \subset A_r$ for each j. If $f^{-1}(A)$ does not have any turns meeting $A(\gamma_4)$ then f can be adjusted so that all crossings meeting $A(\gamma_1) \cup \mathring{J}_{\partial\gamma_1}$ are simple and contracting.*

REMARK. The proof of 5.1 is actually more important than the lemma itself. The method employed will be used in future lemmas (proof of 5.2 and parts of 5.4, 5.5 for instance) which may have slightly different hypothesis but, nonetheless, strong enough to use the same style proof. One purpose of the statement of 5.1 is to serve as an illustration of the method. For comparison we state without proof Lemma 5.2 the proof of which differs only in how 4.1 is applied (see proof of 5.1).

LEMMA 5.2. *Let* γ, γ' *be adjacent* Δ_i-*spirals,* $i \in \{1, 2\}$, *with* γ *nearer* $\mathring{\partial}_i$. *If* $f^{-1}(A)$ *does not have any turns meeting* $A(\gamma')$ *then* f *can be adjusted so that all crossings meeting* $A(\gamma) \cup \mathring{J}_{\partial\gamma}$ *are simple and expanding.*

Proof of 5.1. Let δ_0 be a crossing which meets $A(\gamma_1)$. Let $R_0, \ldots, R_k$ be a sequence of disjoint critical regions satisfying;

   (i) δ_0 meets R_0.
   (ii)   For each $1 \le i \le k$, R_i lies in $C(\gamma_4)$.
   (iii) For each $i < k$, there is a crossing, δ_{i+1}, meeting both R_i and R_{i+1}.

The construction of such a sequence is easy. In fact, starting with δ_0 there are two choices for R_0 (except when δ_0 meets $J_{\partial\gamma_1}$) and once a choice is made R_1, $R_2, \ldots$ are then determined. Now, if $R_0, \ldots, R_k$ satisfies (i)–(iii) then for any $k' < k$, $R_0, \ldots, R_{k'}$ does as well so we consider a maximal k and concentrate on the critical region R_k.

The hypothesis in the lemma imply that any curve which starts on $J_{\partial\gamma_1}$ and extends into $C(\gamma_4)$ must "spiral" outwards until it joins ∂_3. As a result we have the following possibilities for R_k depending on the turns and crossings which R_k meets; (1) R_k meets only δ_k, (2) R_k meets δ_k and δ_0 (here $k = 3$), (3) R_k meets δ' where δ' is either a turn or a crossing which meets $J_{\partial\gamma_1}$ and a critical region not in $C(\gamma_4)$. Eliminate (3) by noting that there is another choice for a maximal sequence and by hypothesis this choice satisfies either (1) or (2). In (2), adjust f so that each of $\delta_0, \ldots, \delta_k$ is simple and contracting so by 4.1.3 none of $R_0, \ldots, R_k$ contains a fixed point. Finally, in case of (1) use 4.1.1 (when $R_k \cap \partial_3 \ne \varnothing$) or 4.1.2 to arrange that δ_k is simple and contracting. Successive applications of 4.1.3 can then be used to get each of $\delta_{k-1}, \ldots, \delta_0$ simple and contracting. $\square$

LEMMA 5.3. *Let* γ *be a* Δ_i-*spiral,* $i \in \{1, 2\}$, *in* $f^{-1}(A)$. *Then* f *can be adjusted so that no turns meet* $A(\gamma) \cup J_{\partial\gamma}$.

Proof. Without loss of generality we take $i = 1$ and $J_{\partial\gamma} \subset A_1$. If γ is not a full Δ_1-spiral then $\gamma = (1, \Delta_1)$ or $(1, 3, \Delta_1)$. Since X is reduced there are no turns meeting $A(\gamma) \cap A_1$ in the former and $A(\gamma) \cap A_3$ in the latter. Also, in the latter case a turn meeting $J_{\partial\gamma}$ must be contained in P_C. Let α be such a turn and R the critical region in P_C meeting α. Let λ_1, λ_2 be the curves in $C(\gamma)$ meeting A_1 adjacent to α. If $\lambda_j \neq \gamma$ for each j or if $\mu(\gamma) \neq 3$ then by 5.2 all crossings meeting $A(\gamma) \cap A_3$ are simple and expanding. It follows that R contains a fixed point (see 4.4) and so we can pull α off of A_1 without increasing the number of fixed points. In the special case $\lambda_1 = \gamma$ with $\mu = 3$ consider the curve, σ, which meets A_3 just outside of γ and extends into P_L. Certainly, $\mu(\sigma) = 1$ and either $\sigma = (3, \Delta_1)$ or it joins α at A_1. In the former, by 4.1.1 we have $\gamma \cap A_3$ simple-expanding and thus a fixed point in R. In the latter we replace $(\sigma, \alpha, \Delta_1)$ by (σ, Δ_1) and make the crossing $\gamma \cap A_3$ simple-expanding. This does not increasing the fixed point count as the resulting critical regions do not contain fixed points.

Now, suppose γ is a full Δ_1-spiral. If α_0 is a minimal turn which meets $A(\gamma) \cup J_{\partial\gamma}$ then either 4.3, 4.4, or the preceding paragraph applies and so we can remove α_0. If no minimal turns occur then each chain of turns "spirals" inwards in $C(\gamma)$ towards ∂_1. Let α denote the maximal turn in an innermost such chain (there may be two) and R the critical region meeting α. By 5.2, all crossings meeting R are simple-expanding and thus R contains a fixed point. In this manner we remove all turns meeting $A(\gamma) \cup J_{\partial\gamma}$. $\qquad\square$

For the next three lemmas we need the following fact concerning curves in $f^{-1}(A)$. The reader is referred to §6 for the validation.

Fact 5.3.1. If γ_1, $\gamma_2 \subset f^{-1}(A)$ are arcs with $\partial\gamma \subset \partial_1 \cup \partial_2$ then γ_t is isotopic (rel ∂P) to either $(\Delta_1, 1, 3, \Delta_1)$ or $(\Delta_1, 3, 4, \Delta_2)$ and γ_1 is isotopic to γ_2.

LEMMA 5.4. *f can be adjusted so that if $f^{-1}(A)$ contains a curve of the form (Δ_3, α, x) where α is a turn then $x = \Delta_i$, $i \in \{1, 2\}$.*

Proof. Without loss of generality we assume $\mu(\alpha_0) = 1$. Suppose $x = 3$ and consider the various possibilities for (Δ_3, α, x, y).

Case 1. If $y = 1$ then by 4.1 and 5.3 there are no turns with $\mu = 3$. Since $f(A_3)$ is a null loop based at a point in P_C we must have that

$f^{-1}(A) \cap A_3 = \varnothing$ (otherwise there is an adjacent pair of points in $f^{-1}(A)$ $\cap A_3$ with the same μ-values, but Lemma 3.4 implies that these points are joined by a turn). But this contradicts the fact that $x = 3$.

Case 2. If $y = 2$ let λ denote the curve meeting A_1 next to α_0 with $\lambda \cap A_1 < \alpha \cap A_1$. If no such λ exists then by 4.1 there are no turns meeting A_3, contradicting $x = 3$ (as in Case 1). Traversing λ in P_C we next reach A_2 (else no turns meeting A_3). Let λ' denote the $(1, 2)$-curve paired with λ and $\delta = (1, 2)$ the other curve (if one exists) adjacent to α.

If $\mu(\lambda) \neq 2$ then by 4.4 there is a fixed point in the critical region meeting α. Pull α off of A_1 without increasing the number of fixed points. If $\mu(\lambda) \neq 2$ and $\mu(\delta) = 2$ then extend δ to one of (δ, Δ_2), $(\delta, 4, 2)$, (δ, Δ_3), or $(\delta, 4, 3, \Delta_3)$. The first two cases give us a Δ_2-spiral and by 4.1, 5.3 their would be no turns meeting A_2. In the latter two cases apply 5.1 to get $\delta \cap A_2$ simple-contracting and using 4.4 detect a fixed point in the critical region meeting α. The case $\mu(\lambda) = 2$, $\mu(\delta) \neq 2$ is handled in the same way which leaves us with $\mu(\lambda) = \mu(\delta) = 2$.

Here λ and δ join to form α_1 with $\alpha \ll \alpha_1$ as indicated in Figure 9. Extend λ_1 to one of (λ_1, Δ_2), $(\lambda_1, 4, 4)$, or $(\lambda_1, 4, 3, \Delta_3)$. Let α_m denote the maximal turn with $\alpha < \alpha_m$. Note that $\mu(\alpha_m) = 2$ or 4. Let ψ denote the curve in P_C meeting A_2 next to λ_1 and extend to one of (ψ, Δ_2), $(\psi, 4, 2)$, or $(\psi, 4, 3, \Delta_3)$ depending on the extension of λ_1. In case of $(\psi, 4, 3, \Delta_3)$ apply 4.4 and 5.1 to detect a fixed point in the critical region meeting α_m which allows us to remove the chain of turns $\alpha, \dots, \alpha_m$. In the other cases we note that their are no turns meeting A_2 besides α_1 and thus, at most one turn, β, meeting A_3. We conclude that β must be minimal for if not there exists a curve isotopic to $(\Delta_1, (3, 1)^{j_1}, 3, 2, (4, 2)^{j_2}, \Delta_2)$. But β minimal

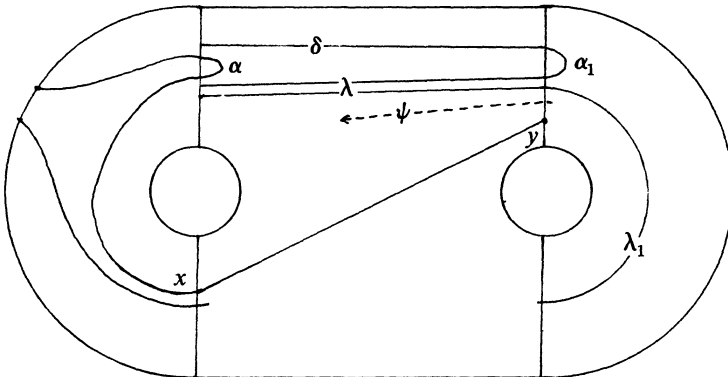FIGURE 9. Case $\mu(\delta) = \mu(\lambda) = 2$

extends to either $(\Delta_1, 1, \beta, 2, \Delta_2)$ or $(\Delta_1, \beta, 1, 2, \Delta_2)$. (Although these curves are not allowed (see 5.3.1), in the case $\mu(\alpha) = 3$ or 4 these would be isotopic to $(\Delta_1, 3, 4, \Delta_2)$ which can occur.) Also, any curve meeting A_3 inside of β must extend to $(\Delta_1, 1, 3, \Delta_1)$. This finishes the case $y = 2$ for if β does not exist we contradict the choice of $x = 3$ and if β exists there must be curves meeting A_3 inside of β contradicting 5.3.1.

Case 3. If $y = 4$ let $\sigma = (\Delta_3, 3, 4)$ denote the curve adjacent to $(\Delta_3, \alpha, 3, 4)$ and consider its possible extensions. In case of (σ, Δ_3) or $(\sigma, 2, 1, \Delta_3)$ use 5.1 to get $\sigma \cap A_3$ simple-contracting and by 4.2 we remove α. With (σ, Δ_2) or $(\sigma, 2, 4)$ there are no turns meeting A_4 inside of σ (5.3). Now, it follows from 5.1 and 4.4 that there is at most one turn meeting A_4 outside of σ. In fact, such a turn must be adjacent to σ. In the present situation, the other curve adjacent to this turn must have the same μ-value as σ which contradicts the fact that X is reduced.

Finally, in case of $(\sigma, 2, 3)$ or $(\sigma, 2, 1, z)$, $z \neq \Delta_3$, consider the curve λ meeting A_1 adjacent to and inside of α. Such a λ exists for if not then $(\Delta_3, \alpha, 3, 4)$ must extend to $(\Delta_3, \alpha, 3, 4, 2, 3, \Delta_1)$ and by 4.1 there would be no turns meeting A_3. Proceed as in case 2 except when $\mu(\lambda) = \mu(\delta) = 2$ and the extension, $(\psi, 4, 3, \Delta_3)$, of ψ is the same curve as the extension of σ. In this instance consider the four critical regions; $R_1 \subset P_L$ meeting σ and α, $R_2 \subset P_C$ meeting α and α_1, $R_3 \subset P_R$ meeting α_1 and $\lambda_1 \cap A_4$, and $R_4 \subset P_C$ meeting $\lambda_1 \cap A_4$ (possibly a turn α_2). Notice that in this situation R_4 also meets the crossing $\sigma \cap A_3$. We finish by adjusting f as follows: Pull the turns α, α_1 off of A_1, A_2 respectively, make the crossing $\sigma \cap A_3$ simple-contracting, and at A_4 pull off the turn α_m or make the crossing $\lambda_1 \cap A_4$ simple-contracting as necessary. In doing so each R_i is replaced by a new critical region R_i' which does not contain a fixed point (4.1, 4.2).   □

LEMMA 5.5. *Suppose* $\alpha \in \Omega(f^{-1}(A))$ *is maximal and contained in* P_C. γ_1, γ_2 *are curves in* P_C *which meet* $A_{\mu(\alpha)}$ *adjacent to* α *with* $\gamma_i = (1, 2)$ *if* $\mu(\alpha) \in \{1, 2\}$ *or* $\gamma_i = (3, 4)$ *if* $\mu(\alpha) \in \{3, 4\}$. *Suppose also that if* $\gamma_i \cap A_{\mu(\gamma_i)} \neq \varnothing$ *then* γ_i *extends* (*from* $A_{\mu(\gamma_i)}$) *to a* Δ_3-*spiral,* γ, *so that* α *is the only turn outside of* $\gamma \cup J_{\partial\gamma}$. *Then we can adjust* f *by removing the turn* α.

REMARK. If no such γ_i occurs outside of α the lemma is still valid as the proof is identical.

Proof. Without loss of generality we assume $\mu(\alpha) = 1$ and $\gamma_2 \cap A_1 < \gamma_1 \cap A_1$. Let R be the critical region in P_C which meets α. We arrange so

that f has a fixed point in R and thus α can be pulled off of A_1 without increasing the number of fixed points.

If $\mu(\gamma_1) = 2$ then by hypothesis γ_1 extends to the Δ_3-spiral γ. Using the method of proof of 5.1 we can arrange that $\gamma_1 \cap A_2$ is simple and contracting and hence R contains a fixed point (4.4).

If $\mu(\gamma_2) = 2$ and extends to γ the curve meeting A_2 just inside of γ extends to a Δ_3-spiral, γ', adjacent to γ with $\mu(\gamma') = 4$. Since no turns meet A_4 outside of γ there must be one inside of γ'. From Lemma 5.3 it then follows that the curve which crosses A_4 just inside of γ' must extend to a Δ_3-spiral, γ'', adjacent to γ'. If $\mu(\gamma'') \neq 3$ then by the proof of 5.1 we can make $\gamma_2 \cap A_2$ simple-contracting otherwise there exists γ''' adjacent to γ'' with $\mu(\gamma''') = 1$. By 5.1 we arrange that $\gamma_2 \cap A_2$ is simple and contracting. Thus, in case $\mu(\gamma_2) = 2$, R contains a fixed point. Finally, if $\mu(\gamma_i) \neq 2$ for each i, then by Lemma 4.4 we can remove α.   □

LEMMA 5.6. *Suppose* $\alpha \in \Omega(f^{-1}(A))$ *is maximal and contained in* $P_L \cup P_R$. *Let* γ_1, γ_2 *denote the curves in* $P_L \cup P_R$ *which meet* $A_{\mu(\alpha)}$ *on either side of* α, $\gamma_1 \cap A_{\mu(\alpha)} < \gamma_2 \cap A_{\mu(\alpha)}$. *Then* f *can be adjusted so that* $\gamma_1 = (\mu(\alpha), \Delta_i)$, $i \in \{1, 2\}$, *and* $\gamma_2 = (\mu(\alpha), \Delta_3)$.

Proof. We assume $\mu(\alpha) = 1$. There are three possibilities; (i) $\gamma_1 = (1, 3)$, $\gamma_2 = (1, \Delta_3)$; (ii) $\gamma_1 = (1, \Delta_1)$, $\gamma_2 = (1, 3)$ or (iii) $\gamma_1 = (1, \Delta_1)$, $\gamma_2 = (1, \Delta_3)$ as desired. Elimination of (i) is almost identical to the proof of 5.4 the reason being that the curves α, γ_1, γ_2 and $\tau = (\Delta_3, 3)$ in this case are virtually the same as the curves $(\Delta_3, \alpha, 3)$ and $(\Delta_3, 3)$ in 5.4.

We consider (ii) with the curve $\tau = (\Delta_1, 3)$ crossing A_3 adjacent to γ_2, $\mu(\tau) = 1$. Let $(\Delta_1, 3, x, y, z)$ denote the extension of τ. The various possibilities for (x, y, z) are handled as follows:

Case 1. $x = 1$. Then γ_2 extends to a full Δ_1-spiral and by Lemma 5.3 we can remove α.

Case 2. $(x, y) = (2, 4)$. By 5.3.1, $z = 3$ and by using 5.4, 5.5, 5.6(i) it follows that no maximal nor minimal turns meet A outside of the curve $(\gamma_2, 2, 4, 3)$. As a result there are no turns outside of $(\gamma_2, 2, 4, 3)$. Hence, the only turns meeting A_3 are on $J_{\partial \tau'}$ where $\tau' = (3, 2, 4, 3) \subset \tau$. But, by 5.4 and 5.6(i), no such turns exist contradicting $A_3 \cap f^{-1}(A) \neq \varnothing$.

Case 3. $(x, y) = (2, \Delta_3)$. Using 5.3-5.5 there are no turns on A_1 outside of α and hence, from 5.4 and 5.6(i), no turns on A_3 outside of τ. Again, contradicting $A_3 \cap f^{-1}(A) \neq \varnothing$.

Case 4. $(x, y) = (4, A_3)$. By 4.2, 4.4, 5.5 there are no turns on A_3.

Case 5. $(x, y) = (4, 2)$. This is similar to Case 2.

Case 6. $(x, y) = (4, \Delta_2)$. If α is not minimal then, by 5.3, 5.4, we must have $\Omega = \{\alpha, \alpha'\}$ with $\alpha' \ll \alpha$. Hence, no turns meeting A_3. If α is minimal then, from 5.3.1, α extends to either $(3, \alpha, 2, 4)$, $(3, \alpha, 2, \Delta_3)$, or $(3, \alpha, 4, \Delta_2)$. In each instance, extend γ_2 and finish as in Case 2.

These six cases cover all allowable possibilities for τ which concludes the proof of 5.6. $\square$

LEMMA 5.7. *f can be adjusted so that given $\alpha \in \Omega(f^{-1}(A))$ extending to the curve (x, α, y) with $x \neq y$ then either $(x, y) = (\Delta_i, \Delta_3)$, $i \in \{1, 2\}$, or $x \in \{1, 3\}$ and $y \in \{2, 4\}$.*

Proof. Let (x, α, y) be given. We assume $\mu(\alpha) = 1$. There are only two possibilities which must be eliminated, namely $(x, \alpha, y) = (\Delta_1, \alpha, 3)$ or $(2, \alpha, 4)$.

The case $(\Delta, \alpha, 3)$ can be taken care of in the same manner as the maximal turn in 5.6(ii) (the settings look almost identical—even the curve $\tau = (\Delta_1, 3)$). In case of $(2, \alpha, 4)$, by using the proof of 5.3, we can remove all turns crossing A_2 or A_4 inside of $(2, \alpha, 4)$ and by 5.2 all crossings inside of $(2, \alpha, 4)$ are simple and contracting. It now follows that the critical region in P_C which meets $(2, \alpha, 4)$ must contain a fixed point and we finish by pulling α off of A_1 without increasing the number of fixed points. $\square$

REMARK. If $\alpha \subset P_C$ then there exists $\tau = (\Delta_1, \Delta_3)$ or $(\Delta_1, 3, 3, \Delta_3)$, $\mu(\tau) = 3$. If $\alpha \subset P_L$ then α extends to $(\Delta_1, (3, 1)^k, 3, \alpha, y)$, $k \geq 0$, $y \in \{2, 4\}$.
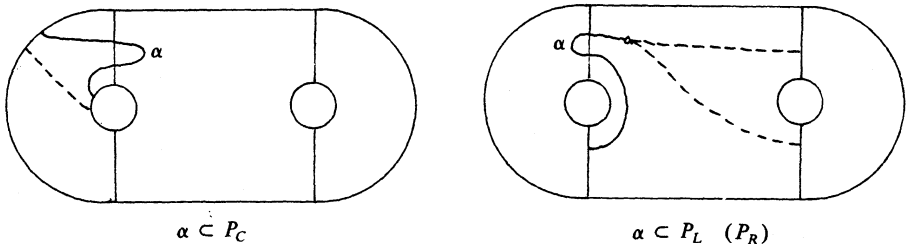

$\alpha \subset P_C$ $\alpha \subset P_L \quad (P_R)$

FIGURE 10. Possibilities for minimal turns $(\mu(\alpha) = 1)$

Given $f\colon P \to P$ let C_f denote the collection of curves obtained from $f^{-1}(A)$ by replacing turns and crossings by simple turns and crossings respectively (as in §4).

PROPOSITION 5.8. *Suppose* $f \simeq F_{(X,Y)}$ *satisfies the conclusions of all previous lemmas and that* (X, Y) *has a type. Then there exist an integer* $0 \le n \le 4$, *a disjoint collection of curves* $\{\beta_i\}_{i=1}^n$ *in* C_f, *curves* $\{\gamma_i\}_{i=1}^n$, *and an isotopy* $H\colon P \times I \to P$ *such that*
  (1) $H = \mathrm{id}$ *on* $(P \times \{0\}) \cup (\partial P \times I) \cup (C_f - \cup\beta_i)$
  (2) β_i *contains exactly one turn,* α_i
  (3) $\partial\gamma_i = \partial\beta_i \subset \partial P \cup A$; $\mathring{\gamma}_i \subset P - A$
  (4) $\mu(\beta_i) = \mu(\beta_j)$ *iff* $i = j$
  (5) $H_1(\beta_i) = \gamma_i$
  (6) $H_1(C_f)$ *contains no turns and all crossings are simple.*

REMARK. If $n = 0$ then $\{\beta_i\} = \varnothing$ in which case $H = \mathrm{id}$.

Proof. We claim that if α, $\beta \in \Omega(C_f)$ with $\mu(\alpha) = \mu(\beta)$ then $\alpha = \beta$. As a result, Ω contains at most four turns. These are the $\{\alpha_i\}$. Also, for each i, there is a unique curve $\beta_i = (x_1^i, \ldots, x_k^i, \alpha_i, y_k^i, \ldots, y_1^i)$ where $x_1^i \ne y_1^i$ and $x_j^i = y_j^i$ for $1 < j \le k$. As $\mu(\alpha_i) \ne \mu(\alpha_j)$ when $i \ne j$ we know that $\beta_i \cap \beta_j = \varnothing$ when $i \ne j$. Set $\gamma_i = (x_1^i, y_1^i)$ then the isotopy H is simply one which takes the collection $\{\beta_i\}$ to the collection $\{\gamma_i\}$ leaving $\partial P \cup (C_f - \cup\beta_i)$ fixed.
  To establish the claim we consider two cases:

Case 1. There exists minimal α_0 contained in P_C. We assume $\mu(\alpha_0)$ is odd then by Lemma 5.7, α_0 extends to the curve $(\Delta_1, \alpha_0, \Delta_3)$ and from 4.1 it is easy to see that A_1 and A_3 each have at most one turn. Suppose now that β is a turn with $\mu(\beta)$ even. Either $\beta \subset P_R$ is maximal or $\beta \subset P_C$ is minimal (in which case the claim is established) or there is a unique turn, β', with $\mu(\beta')$ even and either $\beta \ll \beta'$ or $\beta' \ll \beta$. Consider all such pairs $\{\beta, \beta'\}$ where $\beta \ll \beta'$. If there is at most one such pair then the claim is established.
  Suppose not and let (β, β') be the pair nearest ∂_2 with $\mu(\beta) = 2$. If β is not minimal then there exists $\alpha \ll \beta$ with $\mu(\alpha)$ odd. But, if $\mu(\alpha) = 1$, by 5.5, there are no turns on A_2 outside of β and if $\mu(\alpha) = 3$ there are no turns outside of β' on A_4 contradicting our assumption. So we have that β is minimal and must then extend to $\overline{\beta} = (\Delta_1, 3, \beta, 4, \Delta_2)$. Similarly, β' must be maximal. Now, let $\sigma = (3, 4)$ be the curve which crosses A_4 just

outside of β' and $\tilde{\sigma}$ its maximal extension (see Figure 11). If $(A_3, \sigma) \subset \tilde{\sigma}$ we are done as there are no turns meeting A_2, A_4 outside of β, β' respectively. As $(\Delta_1, \ldots, \Delta_1)$ cannot occur this only leaves $\tilde{\sigma} = (\Delta_1, \sigma, 2, 2, 4, 3, \Delta_3)$. But as $\mu(\sigma) = \mu(\beta) = 2$, by 3.4, there is a curve $\lambda = (\Delta_1, 3, 4, \Delta_2)$ meeting A_3 between σ and $\bar{\beta}$. In fact, all curves between $\bar{\beta}$ are of this form. As a result, $X = Y^{-1}ba--$ with $|Y| \geq 2$ contradicting (X, Y) having a type.

Case 2. All minimal turns are contained in $P_L \cup P_R$. By 5.7, there are at most two minimal turns one in P_L one in P_R. Also, one cannot be maximal so we consider $\alpha_0 \ll \alpha_1$, $\mu(\alpha_0) = 1$, $\mu(\alpha_1) = 3$, α_0 minimal, and β_0 (if it exists) with $\mu(\beta_0) \in \{2, 4\}$. By 5.3 and 5.5 it is not hard to show that there are no other turns on $A_1 \cup A_3$. Likewise when β_0 exists and is not maximal. If β_0 is maximal then by 5.6 we have at most one turn on each of A_2, A_4 and finally by 5.3, 5.5 a single chain cannot have two turns meeting the same A_r.   □
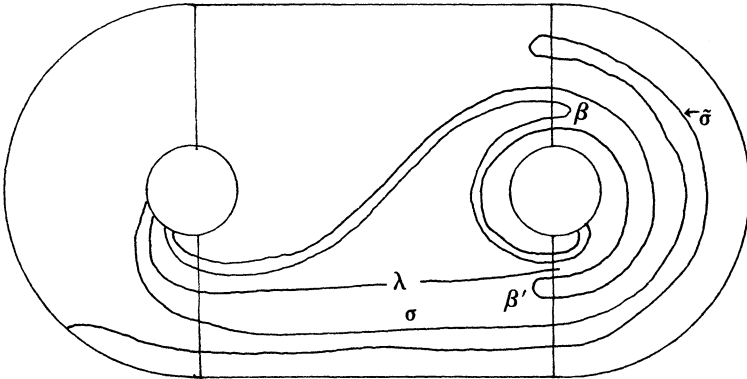

FIGURE 11. "No type"

REMARK. Proposition 5.8 will be applied in the following manner. Given $H_1(C_f)$ consider those arcs in $A - H_1(C_f)$ which satisfy the hypothesis of Lemma 3.4. In order to satisfy the conclusion of 3.4 each of these arcs must be "covered" by a turn. As a result of 5.8, there must be at most one such arc on each A_r. Now, given such a turn in C_f when searching for a candidate for the corresponding γ we need only consider curves (Δ_1, Δ_3), (Δ_2, Δ_3), or (x, y) in P_C such that when γ is replaced by $H_1^{-1}(\gamma)$ the resulting collection has at most one arc as mentioned above on each A_r and none on $A_{\mu(\gamma)}$.

Given γ_i there is possibly two ways to "push" γ_i in order to obtain the curve β_i. If A_r is the first of $A_1, \ldots, A_4$ to be crossed when isotoping γ_i to β_i we say that γ_i *is pushed across* A_r. For example, $\gamma = (3, 4)$ can be pushed across either A_1 or A_2.

6. Isotopy classification of $f^{-1}(A)$. In order to gain an understanding of $f^{-1}(A)$ we begin by considering possibilities, up to isotopy, for this 1-dimensional manifold. The purpose of this section is to classify all such possibilities. In view of the language of this paper we need only consider curves which do not contain any turns.

Let (X, Y) have a type and let f be homotopic to $F_{(X,Y)}$ rel ∂P with $f^{-1}(A)$ being a 1-dimensional proper submanifold of P. Suppose further that Lemmas 3.4 through 3.6 apply to f and so, by Corollary 3.7, $f^{-1}(A)$ does not contain any inessential simple closed curves. For the following let Γ_f denote the collection of curves which is isotopic (rel ∂P) to $f^{-1}(A)$ and contains no turns. The relative position of the points in $\Gamma_f \cap A$ will not be important. As Γ_f depends on the type of (X, Y) we proceed with a case-by-case study. Let $\alpha_0 = (\Delta_1, \Delta_3)$, $\alpha_1 = (\Delta_1, 3, 2, 4, 3, \Delta_)$, and $\alpha_{(-1)} = (\Delta_1, 1, 4, 2, 1, \Delta_3)$. For $n > 1$, α_n is obtained from α_{n-1} by replacing $(3, 2)$ with $(3, 1, 3, 2, 4, 3, 1, 2)$ and for $n < -1$, α_n is obtained from α_{n+1} by replacing $(1, 4)$ with $(1, 3, 1, 4, 2, 1, 3, 4)$.

PROPOSITION 6.1. *If (X, Y) is of type T_1 then there are integers n, m with $0 \leq m < 2|X|$ such that $\Gamma_f = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ where*
   Γ_1 consists of $2|X| - m$ curves isotopic to α_n
   Γ_2 consists of m curves isotopic to α_{n+1}
   Γ_3 consists of simple closed curves isotopic to ∂_2.
Moreover, adjacent curves in Γ_3 have different μ-values.

REMARKS. (1) The μ-values for curves in $\Gamma_1 \cup \Gamma_2$ are determined as $f \simeq F_{(X,Y)}$ rel ∂P. If adjacent curves in Γ_3 have different μ-values these values are determined by considering $\Gamma_3 \cap A_4$ and noting that $f(A_4)$ is a null loop based at a point in P_C. Thus the outermost simple closed curve has the same μ-value as the curve in $\Gamma_1 \cup \Gamma_2$ which crosses A_4 nearest to ∂_2. The rest are paired accordingly. (2) It can be shown that

$$\alpha_n = \begin{cases} \left(\Delta_1, (3, 1)^{n-1}, 3, 2, 4, 3, (1, 2, 4, 3)^{n-1}, \Delta_3 \right) & \text{if } n \geq 1 \\ \left(\Delta_1, (1, 3)^{|n|-1}, 1, 4, 2, 1, (3, 4, 2, 1)^{|n|-1}, \Delta_3 \right) & \text{if } n \leq -1. \end{cases}$$

Proof. For each $g \simeq F_{(X,1)}$ rel ∂P, let $\tau(g)$ denote the minimum cardinality possible for the set T in Lemma 3.3 (among homotopies

between g and $F_{(X,1)}$). Thus, $0 \le \tau(g) < \infty$. The proposition is proved by induction on the integer τ.

For the starting case let f be given so that $\tau(f) = 0$. Let H be a homotopy rel ∂P between f and $F_{(X,1)}$ so that $H_t^{-1}(A)$ is a manifold for all t. Clearly, H_0 satisfies the conclusion of 6.1 as $\Gamma_{F_{(X,1)}} = \Gamma_1$. Since $H_t^{-1}(A)$ is a manifold and H is rel ∂P it follows that each (Δ_1, Δ_3)-curve in H_0 is isotopic to a unique curve in H_1. Hence, Γ_f consists of $2|X|$ (Δ_1, Δ_3)-curves and a finite number of simple closed curves isotopic to ∂_2.

Suppose that δ_1, δ_2 are simple closed curves in $f^{-1}(A)$ (isotopic to ∂_2) bounding an annulus C whose interior misses $f^{-1}(A)$. If $\mu(\delta_1) = \mu(\delta_2)$ we can choose $s \in \{2,4\}$ so that $s \ne \mu(\delta_i)$. As δ_i is isotopic to ∂_2, $\delta_i \cap A_s \ne \varnothing$ and so there exists an arc component, J, of $C \cap A_s$. But Lemma 3.4 says that $\partial J \cap f^{-1}(A) = \varnothing$, a contradiction as $\partial J \subset \delta_1 \cup \delta_2$. Hence, $\mu(\delta_1) \ne \mu(\delta_2)$ which establishes the case $\tau = 0$. Note that $\Gamma_3 = \varnothing$ in this case (Remark (1)).

Now, assume that the proposition holds for any map g with $\tau(g) < k$ and let H be a homotopy rel ∂P between $H_0 = F_{(X,1)}$ and $H_1 = f$ with $T = t_1 < t_2 < \cdots < t_k$. Also, we assume that H is chosen so that $H_z^{-1}(A)$ is proper for some $z \in (t_{k-1}, t_k)$. Then, by the inductive hypothesis, there exist n, m so that $H_z^{-1}(A) = \Gamma_1' \cup \Gamma_2' \cup \Gamma_3'$. Since $H_{t_k}^{-1}(A)$ has only one non-manifold point and adjacent curves in $\Gamma_1' \cup \Gamma_2'$ have different μ-values it follows that at least $2|X| - 1$ of these curves are isotopic (rel ∂P) to corresponding curves in Γ_f. Thus, Γ_1 contains $2|X| - 1$ curves isotopic to either α_n or α_{n+1}. The isotopy type for the remaining arc, γ, depends on the other $2|X| - 1$. If all are isotopic to α_j, $j = n$ or $n + 1$,
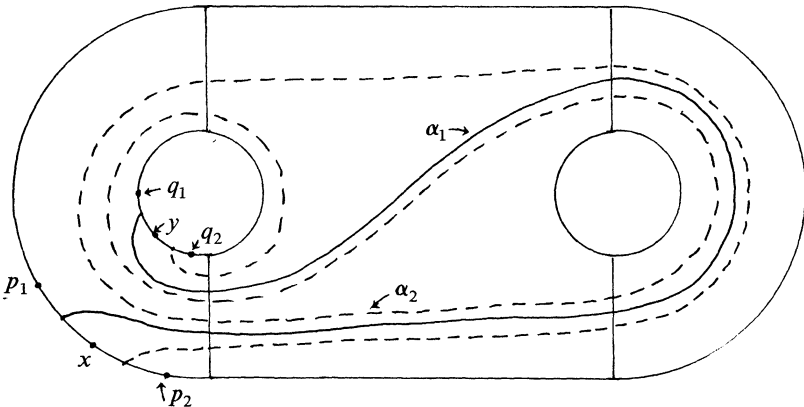

FIGURE 12. α_1 and α_2.

If $\partial \gamma = \{x, y\}$ then γ is either α_1 or α_2.

If $\partial \gamma = \{p_i, q_i\}$, $i \in \{1,2\}$, then $\gamma = \alpha_i$.

then $\gamma = \alpha_{j-1}$, α_j, or α_{j+1}. If both α_n and α_{n+1} curves appear in Γ_f then $\gamma = \alpha_n$ or α_{n+1} (see Figure 12 for an illustration). Thus, Γ_1 and Γ_2 are as desired. As before, all simple closed curves must be isotopic to ∂_2 and the argument given in the $\tau = 0$ case establishes Γ_3.   □

The key to the proof of Proposition 6.1 is that once the inductive hypothesis is applied to $H_z^{-1}(A)$ the isotopy type of all but one proper arc in Γ_f is determined. To find the last one we just consider all possibilities. For types $T_2, \ldots, T_5$ the idea is exactly the same. The difference being that here there are no essential simple closed curves and in the inductive step all but two proper arcs are determined for Γ_f. For the remaining two (which have the same μ-values) there are only two choices to be made. This is simply a result of the classification of pairs (X, Y) into types $T_1, \ldots, T_5$.

With this in mind we obtain Propositions 6.2,...,6.5, the proofs of which are the same in spirit as that of 6.1. Due to the complexity of an explicit formula for curves when (X, Y) is type T_3 only an implicit description is given. Of course, one which contains the information needed for counting fixed points.

PROPOSITION 6.2. *If (X, Y) is of type T_2 then there is an integer m,* $0 \le m \le 2|U|$ *so that $\Gamma_f = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \cup \Gamma_4$ where:*

Γ_1 *consists of $2|\overline{X}| + 4|U| - 2m$ curves isotopic to (Δ_1, Δ_3),*

Γ_2 *consists of m curves isotopic to $(\Delta_1, 3, 1, \Delta_1)$,*

Γ_3 *consists of m curves isotopic to $(\Delta_3, 3, 1, \Delta_3)$,*

Γ_4 *consists of $2|Y|$ curves isotopic to (Δ_2, Δ_3).*

Moreover, as $f \simeq F_{(X,Y)} \mathrm{rel}\, \partial P$ the ν-values of all the curves in Γ_f are determined.

Let $\Gamma^0 = F_{(X,Y)}^{-1}(A)$ and for certain integers n define Γ^n inductively as follows:

For $n > 0$ we assume for now that Γ^{n-1} is defined and satisfies alll conditions mentioned below. Let γ_1 be either the (Δ_1, Δ_3)-curve in Γ^{n-1} nearest (in P_L) to A_3 or, if no such curve exists, the $(3, 2)$-curve nearest (in P_C) to A_1. Similarly, γ_2 is either the (Δ_2, Δ_3)-curve nearest A_2 or the $(3, 2)$-curve nearest A_4. Let I be an arc with interior in $P - \Gamma^{n-1}$ which joins γ_1 to γ_2 and has minimal intersection with A. We remark that I is unique up to isotopy and must contain a $(3, 2)$-curve. Let I_1, I_2 be arcs
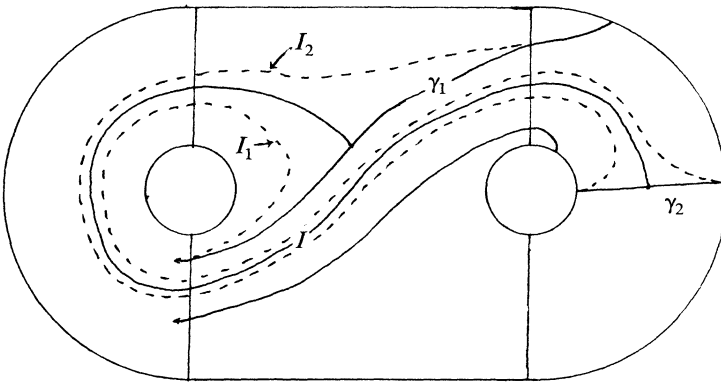
FIGURE 13

with interior missing $I \cup \Gamma^{n-1}$ chosen close to $I \cup \gamma_1 \cup \gamma_2$ and with $\partial(I_1 \cup I_2) = \partial(\gamma_1 \cup \gamma_2)$ (see Figure 13). If $\mu(\gamma_1) = \mu(\gamma_2)$ (see μ-values below) we set $\Gamma^n = (\Gamma^{n-1} - (\gamma_1 \cup \gamma_2)) \cup (I_1 \cup I_2)$ otherwise we do not define Γ^n.

For $n < 0$ we assume that Γ^{n+1} is given and define γ_1, γ_2, I, I_1, I_2 in the same manner except that A_3 and A_1 (respectively A_4 and A_2) are interchanged and $(3,2)$ is replaced by $(1,4)$. Γ^n is then defined analogously.

REMARK. It is easy to verify that the assumptions on Γ^{n-1} are valid. Also, if $\gamma \in \Gamma^n$ is not equal to (Δ_j, Δ_3) then γ can be expressed in the form (β_1, x, y, β_2) where one of β_1, β_2 is a Δ_3-spiral and the other is a Δ_i-spiral, $i \in \{1,2\}$, and (x, y) is either $(3,2)$ or $(1,4)$.

We assign μ-values to the curves in Γ^n as follows. Let $C_1, \ldots, C_{2|X|}$ denote the μ-values of the (Δ_1, Δ_3)-curves in Γ^0 (from $F_{(X,Y)}$) given in order starting with the curve nearest A_3. The μ-values of a (Δ_j, Δ_3)-curve in Γ^n is the same as the corresponding curve in Γ_0. For $n > 0$, the μ-values for the $(3,2)$-curves in Γ^n are (as they meet A_3 starting from ∂_1):

$$C_1, \ldots, C_n, *, C_n, \ldots, C_1 \quad \text{if } 0 < n \le 2|W|;$$

$$C_1, \ldots, C_n, *, C_n, \ldots, C_{(n-2|W|+1)} \quad \text{if } 2|W| < n \le 2|X|;$$

$$C_{(n-2|X|+1)}, \ldots, C_n, *, C_n, \ldots, C_{(n-2|W|+1)} \quad \text{if } 2|X| < n.$$

For $n > 2|X|$ we set $C_n = C_{n-2|W|}$. Also, $*$ denotes the placement of $I \cap A_3$. For $n < 0$ the μ-values for $(1,4)$-curves are obtained as above with A_1 replacing A_3.

It is an easy exercise in induction to check that this assignment yields the same μ-values for Γ^n as those induced by Γ^{n-1} (or Γ^{n+1} when $n < 0$) where $\mu(\gamma_1) = \mu(\gamma_2) = \mu(I_1) = \mu(I_2)$. Also, as the construction of the Γ^n's is modeled on the proof of 6.1 we have,

PROPOSITION 6.3. *If (X, Y) is of type T_3 then $\Gamma_f = \Gamma^n$ for some integer* n.

REMARK. $n > 2|W|$ implies that $\overline{Y} = 1$; $n > 2|X|$ implies that $\overline{\overline{X}} = 1$; and $n < 0$ implies that $X = V^{k_1}$, $Y = V^{k_2}$, $0 < k_2 \le k_1$, for some V.

PROPOSITION 6.4. *If (X, Y) is of type T_4 then there is an integer* $0 \le m \le 2|W|$ *so that $\Gamma_f = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \cup \Gamma_4$ where*

Γ_1 *consists of $2|\overline{X}| + 2|W| - m$ curves isotopic to (Δ_1, Δ_3)*

Γ_2 *consists of m curves isotopic to $(\Delta_1, 3, 4, \Delta_2)$*

Γ_3 *consists of m curves isotopic to $(\Delta_3, 3, 4, \Delta_3)$*

Γ_4 *consists of $2|\overline{Y}| + 2|W| - m$ curves isotopic to (Δ_2, Δ_3).*

Moreover, $f \simeq F_{(X,Y)} \operatorname{rel} \partial P$ determines all μ-values.

PROPOSITION 6.5. *If (X, Y) is of type T_5 then Γ_f is as in Proposition 6.4 except that $|W| = 1$, $|\overline{Y}| = 0$, and $|l_2\overline{X}|$ replaces $|\overline{X}|$.*

7. Calculation of $MF(F_{(X,Y)})$. The proof of Theorem 1.1 will now be concluded. Let (X, Y), having a type, and $f \simeq F_{(X,Y)} \operatorname{rel} \partial P$, satisfying the conclusions of all previously stated Lemmas, be given. Depending on the type for (X, Y) proceed to the appropriate one of (7.1)–(7.5). In each case the proof goes as follows. First, a description of $f^{-1}(A)$ is given (in terms of Γ_f) together with necessary conditions on the pair (X, Y) for this to occur. Then using Lemma 2.1, we show that $\#\operatorname{Fix}(f) \ge M(X, Y)$. Equality is established by noting that each critical region which is not mentioned in the proof need not have a fixed point. The verification of this last statement is, in each of (7.1)–(7.5), a simple application of Lemma 2.1 and is (in general) omitted.

(7.1) (X, Y) has type T_1. Let $\Gamma_f = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ be as given in Proposition 6.1. A detailed analysis is given below in the case $n > 0$ and brief summaries follow for $n \le 0$.

The first thing to do is to find all possibilities for the $\{\gamma_i\}$ of Proposition 5.8. First note that a $(2, 4)$-curve in Γ_3 is not a candidate as this would mean that C_f contains the curve $(2, 3, 3, 4)$ contradicting

Lemma 5.7. Similarly, $(1, 3)$-curves are eliminated. By the remark after 5.8 we cannot push $(1, 2)$, $(3, 4)$, and $(3, 2)$-curves across A_3, A_2, and A_4 respectively. As a result the only possibility for forming a turn in C_f is by pushing a $(3, 2)$-curve across A_1 and thus that $\Omega(C_f)$ consists of a single chain of four turns. We set $\gamma_1 = (3, 2)$-curve which crosses A_3 nearest to ∂_1 and γ_2, γ_3, γ_4 successively parallel $(3, 2)$-curves. Then $\beta_1 = (3, 1, 3, \lambda, 3, 1, 2)$ where $\lambda = (2, 4, 4, 2)$ or $(4, 2, 2, 4)$ and β_i, $1 < i \leq 4$, are determined accordingly. It follows that $\mu(\gamma_1)$ and $\mu(\gamma_2)$ are even, $\mu(\gamma_3) = 3$, and $\mu(\gamma_4) = 1$ which implies that $X = X_1 b^{\pm 1} a X_2$ or $a X_1 b^{\pm 1}$.

Since each curve in $\Gamma_1 \cup \Gamma_2$ contains exactly one $(3, 2)$-curve there are $\Phi_a(X)$ such curves having $\mu = 1$, one being γ_4 and the rest appearing as in C_f. Each one of the latter meets a distinct critical region contained in P_C. This case is finished by showing that each such region contains a fixed point of f.

Let R denote one of these regions with $\beta = (3, 2) \subset \partial R$, $\mu(\beta) = 1$. There are two possibilities to consider: (i) $\partial R \cap f^{-1}(A)$ consists of two $(3, 2)$-curves, β and δ with $\mu(\delta) \neq 1$ and (ii) ∂R contains the maximal turn, α, of $\Omega(f^{-1}(A))$ with $\mu(\alpha) = 2$. In (i) use Lemmas 5.1, 5.2 to show that if $\mu(\delta) = 3$ then the crossing $\delta \cap A_3$ is simple-expanding and if $\mu(\delta) = 2$ simple-contracting. To detect a fixed point in R set up a product structure for P_C so that $\mathscr{S} = \{ p, q \}$ with $p = \beta \cap A_3$ and $q \in \partial \delta$. From the preceding discussion it follows that p, q have opposite labels and so R contains a fixed point. In (ii) use the same product structure so that $\mathscr{S}$ consists of $\alpha \cap A_2$, $p = \beta \cap A_3$, and a point q in A_4. Since $f(q) \subset A_2$ both p and q are labeled the same and so $\mathscr{S}' = \alpha \cap A_2 \neq \varnothing$. Hence $f \mid R$ has a fixed point.

To summarize $n \leq 0$, first consider the case when $n = 0$ and $m > 0$. Here $\{ \gamma_i \}$ are taken from Γ_1 starting with the (Δ_1, Δ_3)-curve nearest A_3. (If $m = 2|X| - 2$ then we take γ_3, $\gamma_4 = (3, 2)$-curves from Γ_2.) Here, we must have $X = X_1 b^{\pm} a X_2$ and the $\Phi_a(X) - 1$ fixed points are found in critical regions meeting either $(3, 2)$ or (Δ_1, Δ_3)-curves with $\mu = 1$ (except γ_4). When $n < 0$ the argument is the same as above with the roles of 1 and 3 (2 and 4) interchanged. In this case, $X = X_1 ab^{\pm 1} X_2$ or $b^{\pm 1} X_1 a$.

Finally, the case $n = m = 0$ has two possibilities, either $\Omega(f^{-1}(A))$ has a chain consisting of three turns or there is at most two turns each being minimal. In the former, the detection of $\Phi_a(X) - 1$ fixed points $(X = X_1 ab^{\pm 1}$ or $b^{\pm 1} a X_1)$ is done as in the previous cases. In the latter, there are no conditions on X and the number of minimal turns can be taken to be $\rho_a(X)$. It is easy to find $\Phi_a(X) - \rho_a(X)$ fixed points in P_L and when $\rho_a(X) \neq 1$ there is an additional fixed point found in the only

critical region in P_C. Notice that since $\rho_b(1) = 0$ that f has at least $\Phi_a(X) + \sigma(X, 1)$ fixed points.

Finally, all other critical regions in the argument above need not contain fixed points and so we have established that $MF(F_{(X,1)}) = M(X, 1)$.

(7.2) (X, Y) has type T_2. Let Γ_f be given as in Proposition 6.2.

Case: $m = 2|U|$, $U \neq 1$. By 5.7, 5.8, $\Omega(f^{-1}(A))$ only contains minimal turns and since we must have turns at A_1 and A_3 resulting from the outermost curves in Γ_1 it follows that $\overline{X}_b = a$, $\overline{X}_e = a$, and $|\overline{X}| > 1$ (in case $\overline{X} = a$ we can remove one of the turns without increasing the number of fixed points—hence a different case is used). This is precisely the case when $\rho_a(X) = 2$ in the formula for $M(X, Y)$. To count fixed points first arrange that all turns are simple-contracting, crossings in Γ_2 are simple-expanding, and crossings in Γ_3 are simple-contracting. Let β denote the curve in Γ_3 which is nearest ∂_2. We count fixed points in the components of $P - \beta$.

To the right: The number of fixed points depends on the μ-values of the two curves δ_1, δ_2 in Γ_4 adjacent to A_2, A_4 respectively and on $\mu(\beta)$. It is easy to verify that we must have $\Phi_b(Y) + \sigma(X, Y) + \Phi_a(U_b)$ such fixed points. Notice that the value of $\sigma(X, Y)$ only depends on $\mu(\beta)$, $\mu(\delta_1)$ and $\mu(\delta_2)$.

To the left: In P_L we detect $\Phi_a(\overline{X}) - 2$ fixed points. In P_C a fixed point in the critical region meeting the turns at A_1 and A_3 as well as one fixed point for each critical region bounded by a pair of curves from Γ_3; one of which has an odd μ-value the other even. Totaling, there are $\Phi_a(\overline{X}) - 1 + \Phi_{\{ab,ab^{-1},a^{-1}b,a^{-1}b^{-1}\}}(U)$ such regions.

Case: $0 < m < 2|U|$. Here no turns arise from curves in Γ_2 and Γ_3 and so the count "to the right" is as in the first case. Let α_1, α_2 denote the curves in Γ_1 adjacent to A_1, A_3 respectively. Clearly, $\mu(\alpha_1) = \mu(\alpha_2)$ and so there are only two ways to form turns at A_1 and A_3. The first is if $\gamma_1 = \alpha_1$ is pushed across A_3 ($\mu(\alpha_1) = 3$) and γ_2 is the (Δ_1, Δ_3)-curve adjacent to γ_1 with $\mu(\gamma_2) = 1$. This corresponds to the condition that aa appears in $\overline{X}_e U^{-1}$. The second occurs when $\gamma_1 = \alpha_2$ ($\mu(\gamma_1) = 1$) and $\mu(\gamma_2) = 3$ which corresponds to the condition that aa appears in $U\overline{X}_b$.

To calculate the number of fixed points we concentrate on the case in which $\gamma_1 = \alpha_1$. The other is handled in the same manner. Following the procedure for $m = 2|U|$ we need to find fixed points "to the left" of β.

This breaks into two cases depending on the correspondence between the curves γ_1, γ_2 and the word $\overline{X}_e U^{-1}$ (equivalently—the value of m).

When $0 < m < 2|U| - 1$, aa appears in U^{-1} and we write $U^{-1} = U_1 aa U_2$ where U_1 corresponds to Γ_1 curves and U_2 corresponds to $\Gamma_2 \cup \Gamma_3$ curves.

For each $\delta \in \Gamma_1 - \{\gamma_2\}$ with $\mu(\delta) = 1$ we find a critical region in P_L containing a fixed point. In P_C, fixed points are located between Γ_3 curves as in the first case. Hence, there are

$$\Phi_a\big(a^{-1}U_1^{-1}XU_1\big) + \Phi_{\{ab,ab^{-1},a^{-1}b,a^{-1}b^{-1}\}}\big(U_2^{-1}a^{-1}\big)$$

fixed points "to the left" of β. This equals

$$2\Phi_a\big(a^{-1}U_1^{-1}\big) - 1 + \Phi_{\{ab,ab^{-1},a^{-1}b,a^{-1}b^{-1}\}}\big(U_2^{-1}a^{-1}\big) + \Phi_a(\overline{X})$$

and by the inequality

$$2\Phi_a\big(a^{-1}U_1^{-1}\big) \geq \Phi_{\{ab,ab^{-1},a^{-1}b,a^{-1}b^{-1}\}}\big(a^{-1}U_1^{-1}\big) + \Phi_a(U_e) + 1$$

(equality iff $\Phi_{aa}(U_1^{-1}) = 0$. Note that $U_e = (a^{-1}U_1^{-1})_e$) we have

$$\Phi_{\{ab,ab^{-1},a^{-1}b,a^{-1}b^{-1}\}}\big(U_2^{-1}a^{-1}a^{-1}U_1^{-1}\big) + \Phi_a(U_e) + \Phi_a(\overline{X})$$

fixed points. When $m = 2|U| - 1$ we have $\overline{X}_e U_e^{-1} = aa$ and counting as we did above there are

$$\Phi_{\{ab,ab^{-1},a^{-1}b,a^{-1}b^{-1}\}}(U) + \Phi_a(\overline{X}) - 1$$

fixed points "to the left" of β. Notice that when combined with the fixed points "to the right" these calculations match the formula for $M(X, Y)$ in the respective cases $\Phi_{aa}(U) > 0$, $\overline{X}_e U_e^{-1} = aa$.

Case: $m = 0$. For this Γ_f there are no conditions on X and Y. Here $\Omega(f^{-1}(A))$ has $\rho_a(X) + \rho_a(Y)$ minimal turns (depending on U_b, Y_b, and Y_e) which leads to $\Phi_a(X) + \Phi_b(Y) + \sigma(X, Y)$ fixed points.

Finally, given the pair (X, Y) more than one of the above mentioned conditions may be satisfied. The one which yields the fewest number of fixed points will give the value of $MF(F_{(X,Y)})$. This is achieved by the following set of priorities: First, $U_e \overline{X}_b = aa$ or $\overline{X}_e U_e^{-1} = aa$ or $\rho_a(\overline{X}) = 2$; second, $\Phi_{aa}(U) > 0$ (here write $U = U_1(aa)^{\pm 1}U_2$ where $\Phi_{aa}(U_2) = 0$ to obtain the formula); third, all others. The validity of the formula now follows using the inequality

$$2\Phi_a(U) \geq \Phi_{\{ab,ab^{-1},a^{-1}b,a^{-1}b^{-1}\}}(U) + \Phi_a(U_b).$$

(7.3) (X, Y) has type T_3. Let $\Gamma_f = \Gamma^n$ as given in Proposition 6.3. In order to verify the formula in this case the following lemma will be used to reduce the number of possibilities that may occur and also as an aid in counting fixed points.

LEMMA 7.3. *Suppose $f^{-1}(A)$ has a critical region, R, in P_C meeting a $(3, 2)$-curve $((1, 4)$-curve$)$, β, and a turn, α, at either A_1 or A_4 $(A_2$ or $A_3)$ with $\mu(\alpha) \neq \mu(\beta)$. Then $f \mid R$ has a fixed point. Moreover, if α is not minimal one can apply the lemmas of the preceding sections to replace Γ^n with $\Gamma^{n'}$ where $|n'| < |n|$.*

Proof. Apply 5.1, 5.2 where necessary so that crossings are simple and contracting/expanding as desired. A fixed point can now be detected in R. By pulling the turn α off of $A_{\mu(\alpha)}$ we do not increase the number of fixed points and by applying methods of previous sections we obtain $\Gamma_{n'}$. □

Below is a list of all the possibilities (i.e. the $\{\gamma_i\}$ of Proposition 5.8) for obtaining $f^{-1}(A)$ from Γ_f. Included in each are the corresponding conditions on the pair (X, Y) and a count of the number of fixed points which arise. As this count is roughly the same in each instance the details are only given in (2) and (6). In (2)–(8) this count is given in the form $\Phi_a(X) + \Phi_b(Y) + \lambda_3$ where $\lambda_3 \in \{-2, -1, +1\}$.

For $n = 0$;
(1) All turns are minimal with no conditions on (X, Y). As in 7.2 we have $\Phi_a(X) + \Phi_b(Y) + \sigma(X, Y)$ fixed points.

For $0 < n < 2|W|$;
(2) $\gamma_1 = (\Delta_1, \Delta_3)$-curve nearest A_3, $\mu(\gamma_1) = 2$. γ_2 next to γ_1, $\mu(\gamma_2) = 3$.

CONDITION: $b^{-1}a$ appears in $W\overline{X}_b$.
FIXED POINT COUNT: Set $W\overline{X}_b = W_1 b^{-1} a W_2$. Fixed points are found in the following critical regions; in P_L those bounded by a (Δ_1, Δ_3)-curve with $\mu = 1$, in P_R those bounded by a (Δ_2, Δ_3)-curve with $\mu = 4$, and in P_C those bounded by $(3, 2)$-curves with $\mu = 1$ if the curve meets A_3 on ∂_1 side of $A_3 - I$ and $\mu = 4$ on the ∂_3 side. From Proposition 6.3 and Lemma 7.3 there are respectively, $\Phi_a(W_2 \overline{X}_b^{-1} \overline{X})$, $\Phi_b(W_2 \overline{X}_b^{-1})$ (except when $W_2 = 1$), and $\Phi_a(W_1 b^{-1}) + \Phi_b(W_1 b^{-1})$ such regions. Combine to obtain

$$\Phi_a\left(W_1 b^{-1} a W_2 \overline{X}_b^{-1} \overline{X}\right) - 1 + \Phi_b\left(W_1 b^{-1} a W_2 X_b^{-1}\right) = \Phi_a(X) + \Phi_b(Y) - 1$$

fixed points. Hence $\lambda_3 = -1$.

(3) $\gamma_1 = (\Delta_2, \Delta_3)$-curve nearest A_2, $\mu(\gamma_1) = 3$. Either γ_2 is next to γ_1, $\mu = 2$ or $\gamma_2 = (3,2)$-curve nearest A_4, $\mu = 2$, and γ_3 is next to γ_2, $\mu = 4$.
CONDITION: $a^{-1}b$ appears in $W\overline{Y}_b$ or $\overline{Y} = 1$ and $W = b$----a^{-1}.
COUNT: $\lambda_3 = -1$ (exactly as in (2)).

For $2|W| \le n < 2|X|$;
(4) $n = 2|W|$, $\gamma_1 = (\Delta_1, \Delta_3)$-curve, $\mu = 3$, $\gamma_2 = (\Delta_2, \Delta_3)$-curve, $\mu = 2$.
CONDITION: $\overline{X}_b = a$ and $\overline{Y}_b = b$.
COUNT: $\lambda_3 = -2$ (exactly as in (2)).

(5) $\gamma_1 = (\Delta_1, \Delta_3)$-curve, $\mu = 3$; γ_2, γ_3 are $(3,2)$-curves, $\mu = 2$, 4 respectively.
CONDITION: $\overline{Y} = 1$, $\overline{\overline{X}}_b = a$, and $(\overline{W}^{-1}W)_b = b$.
COUNT: $\lambda_3 = -2$ (same as (2) except that the $(3,2)$-curves "below" I correspond to a permutation of W which contributes $\Phi_b(W) - 1$ fixed points).

For $n > 2|W|$;
(6) $\gamma_1 = (\Delta_1, \Delta_3)$ or $(3,2)$-curve nearest A_3 (resp. A_1), $\mu = 4$; γ_2, γ_3, γ_4 following with μ-values 2, 3, 1 respectively.
CONDITION: $\overline{Y} = 1$, and $b^{-1}a$ appears in $W^{N-1}\overline{W}\overline{\overline{X}}_b$ or $\overline{\overline{X}} = 1$ and $W^N\overline{W} = a$----b^{-1}.
COUNT: Consider the case $n \ge 2|X|$ (hence $\overline{\overline{X}} = 1$). $n < 2|X|$ is calculated in the same manner. From 6.3, the μ-values of $(3,2)$-curves in Γ^n are $C_{\bar{n}+1}, \ldots, C_n, *, C_n, \ldots, C_{n-2|W|+1}$ where $\bar{n} = n - 2|X|$. By hypothesis, $C_{\bar{n}+1}, \ldots, C_{\bar{n}+4}$ are respectively 4, 2, 3, 1. Obtaining $f^{-1}(A)$ from Γ^n rearranges $(3,2)$-curves to give

$$C_{\bar{n}+5}, \ldots, C_n, C_{\bar{n}+1}, C_{\bar{n}+2}, *, C_{\bar{n}+2}, C_{\bar{n}+1}, C_n, \ldots, C_{n-2|W|+1}.$$

Note that those before $*$ correspond to $\tilde{X}b^{-1}$ where $b^{-1}a\tilde{X}$ is a permutation of X (length $(n - 2|X|)/2$ prefix is moved to end); the latter to $\tilde{W}b^{-1}$ where $\tilde{W}$ is the same permutation of W. By Lemma 7.3, $C_{n-2|W|+1} = 4$ and so $\tilde{W}_b = b^{-1}$. Counting fixed points as in (2) we have $\Phi_a(\tilde{X})$ above $*$ and $\Phi_b(\tilde{W}b^{-1}) - 1$ below (the region meeting $C_{n-2|W|+1}$ need not have a fixed point). Combine to obtain $\lambda_3 = -1$.

(7) $\gamma_1 = (3,2)$-curve nearest A_4, $\mu = 1$; γ_2, γ_3, γ_4 following with μ-values 3, 2, 4 respectively.

CONDITION: $\overline{Y} = 1$, and $a^{-1}b$ appears in W or $W = b----a^{-1}$.

COUNT: $\lambda_3 = -1$ (exactly as in (6)). *For $n < 0$; ($\overline{Y} = 1$, $\overline{\overline{X}} = 1$, and $X_e = Y_e$)*

(8) Situations symmetrical to (2), (3), (6), and (7) occur. The fact that $X^{k_1} = Y^{k_2}$ for some k_1, k_2 eliminates (4) and (5).

CONDITIONS: ab^{-1} appears in $W^N\overline{W} = X$, $X = b^{-1}---a$, ba^{-1} appears in W, or $W = a^{-1}---b$ are the corresponding conditions.

COUNT: $\lambda_3 = -1$.

Now, the value of $MF(F_{(X,Y)})$ is just the smallest possible count for those of (1)–(8) which apply to the pair (X, Y). Hence, the value of λ_3 in the statement of the formula should be the minimum over cases (2)–(8). This is the case except for the following three discrepancies. First, when $W = b---a^{-1}$ or $W^N\overline{W} = a---b^{-1}$ we should have $\lambda_3 = -1$ not 1. But as $\sigma = -1$ in this case the value of $M(X, Y)$ is correct as stated. Secondly, in (6) the condition that $b^{-1}a$ appears in $W^{N-1}\overline{W}$ is not accounted for in the statement of $M(X, Y)$ but in this case either $b^{-1}a$ appears in W or $W = a---b^{-1}$. Hence, either $\Phi_{a^{-1}b}(W) > 0$ or $\sigma = -1$ as was needed. Similarly, if the conditions for (8) are satisfied then it is easy to check that either $\sigma = -1$ or $\Phi_{a^{-1}b}(W) > 0$ as well.

(7.4) (X, Y) has type T_4. Let Γ_f be as given in Proposition 6.4 and consider cases depending on the value of m;

For $m = 0$;

(1) All turns are minimal with no conditions on (X, Y). There are $\Phi_a(X) + \Phi_b(Y) + \sigma(X, Y)$ fixed points in this case.

For $0 < m < 2|W|$;

(2) $\gamma_1 = (\Delta_1, \Delta_3)$-curves nearest A_3, $\mu = 4$; γ_2 next with $\mu = 3$.

CONDITION: ba appears in $W\overline{X}_b$.

COUNT: In case $m < 2|W| - 1$ write $W = W_1baW_2$. Count fixed points as in (7.1)–(7.3) to obtain;

$$\Phi_a(W_2\overline{X}) - \begin{cases} 1 & \text{if } \overline{X}_e = a \\ 0 & \text{otherwise} \end{cases} \quad \text{fixed points in } P_L,$$

$$\Phi_b(\overline{Y}W_2^{-1}) - \begin{cases} 1 & \text{if } \overline{Y}_b = b \\ 0 & \text{otherwise} \end{cases} \quad \text{fixed points in } P_R,$$

and one fixed point for each critical region in P_C which meets two $(3, 4)$-curves (which extend to $(\Delta_1, 3, 4, \Delta_2)$); one with $\mu \in \{1, 2\}$ and the other in $\{3, 4\}$. There are $\Phi_{\{aa,bb,ab^{-1},ba^{-1},a^{-1}b,b^{-1}a\}}(W_1ba)$ such regions. Also, there is a fixed point in the critical region in P_C bounded by the upper $(3, 4)$-curve except when exactly one of $\overline{X}_e = a$, $\overline{Y}_b = b$, $W_b = a$ or b^{-1} holds true. Using the fact that

$$\Phi_{\{aa,bb,ab^{-1},ba^{-1},a^{-1}b,b^{-1}a\}}(W_1ba) + \Phi_{\{ab,ba\}}(W_1ba)$$

$$= \Phi_a(W_1ba) + \Phi_b(W_1ba) - 1$$

and combining the three "conditionals" into λ_4 ($X' = X$; $Y' = y$) we obtain $\Phi_a(X) + \Phi_b(Y) + \lambda_4 - \Phi_{\{ab,ba\}}(W_1ba) - 1$ fixed points. This corresponds to (II) in formula for $M(X, Y)$.

In case $m = 2|W| - 1$, write $W = W_1b$ and count as above to obtain $\Phi_a(X) + \Phi_b(Y) + \lambda_4 - \Phi_{\{ab,ba\}}(W) - 2$ fixed points (see (I) in formula).

(3) $\gamma_1 = (\Delta_2, \Delta_3)$-curve nearest A_4, $\mu = 3$; γ_2 next with $\mu = 4$.
CONDITION: ba appears in $\overline{Y}_eW^{-1}$.
COUNT: for $m < 2|W| - 1$ write $W = W_1a^{-1}b^{-1}W_2$ to obtain (II) in statement (exactly as in (2) above). For $m = 2|W| - 1$ we obtain (I).

For $m = 2|W|$;
(4) $\gamma_1 = (\Delta_1, \Delta_3)$-curve nearest A_3, $\mu = 3$; $\gamma_2 = (\Delta_2, \Delta_3)$-curve nearest A_4, $\mu = 4$.
CONDITION: $\overline{X}_b = a$ and $\overline{Y}_e = b$.
COUNT: As in (2) to obtain (I).

REMARKS. (i) The reason for X' and Y' in the statement of the formula is to avoid counting problems in case $\overline{X} = a$ or $\overline{Y} = b$.

(ii) If (X, Y) satisfies conditions for both (I) and (II) the smaller must be chosen for $MF(F_{(X,Y)})$. The value of λ_4 may be different in the two cases but, nevertheless, the value in (I) is less than or equal to the value in (II). Similarly, if $\Phi_{ba}(W) > 1$ then the last appearance of either ba or $a^{-1}b^{-1}$ yields the fewest number of fixed points. Hence, the condition $\Phi_{ba}(W_2) = 0$ in the statement of (II).

(7.5) (X, Y) has type T_5. Let Γ_f be as given in Proposition 6.5. The flavor of this part will be slightly different than the preceding in that knowledge of l_1l_2 will restrict the number of possibilities for $\{\gamma_i\}$ which could have otherwise occurred. We consider cases depending on the value of m.

Case $m = 0$: As usual there are no conditions on (X, Y) and there are $\Phi_a(X) + \Phi_b(Y) + \sigma(X, Y)$ fixed points.

Case $m = 1$: First consider $\gamma_1 = (\Delta_2, \Delta_3)$-curve. If $\mu = 1$ then $l_1 = a$ and γ_1 forms a non-minimal turn at A_1. We must have $\mu(\gamma_2) = 2$ but the only choice for γ_2 is the (Δ_1, Δ_3)-curve nearest A_3 which has $\mu = 1$. If $\mu = 3$ then $l_1 = a^{-1}$ and γ_1 forms a turn at A_3. Here we must have $\gamma_2 = (\Delta_1, \Delta_3)$-curve nearest A_1 with $\mu(\gamma_2) = 4$ but this would contradict the fact that $X_e = b^{-1}$. As a result, we are left with the choice; $\gamma_1 = (\Delta_1, \Delta_3)$-curve nearest A_3, $\mu = 4$ and γ_2 next with $\mu = 3$.

CONDITION: $l_1 = b$ and hence $l_2 = a$

COUNT: If $\overline{X} = 1$ there is exactly one fixed point (in P_C). If $\overline{X} \neq 1$ then $\overline{X}_e = a$ and there are $\Phi_a(\overline{X}) - 1$ fixed points (all in P_L). In all, $\Phi_a(X) - 2\rho_a(\overline{X}_e)$.

Case $m = 2$: The possibilities for γ_1 are; (i) $(3, 4)$-curve nearest ∂_1 with $\mu = 4$, $l_1 = b^{-1}$, (ii) (Δ_1, Δ_3)-curve nearest A_1, $\mu = 4$, $l_2 = b$, (iii) (Δ_1, Δ_3)-curve nearest A_3, $\mu = 3$, $l_2 = a$, and (iv) (Δ_1, Δ_3)-curve nearest A_3, $\mu = 2$, $l_2 = b$. In (i), $l_2 = a^{-1}$ and so there is no turn at A_3. In (ii), there is also no turn at A_3 and in (iii) there is no turn at A_4. Hence, (i)–(iii) cannot possibly occur which only leaves (iv). In this case, γ_2 is next to γ_1 and γ_3 is either the next (Δ_1, Δ_3) or, if no more exist, the $(3, 4)$-curve nearest ∂_1, $\mu(\gamma_3) = 3$.

CONDITION: $l_2 = b$ and either $\overline{X} = 1$ or $\overline{X}_b = a$.

COUNT: The critical region in P_C meeting the maximal turn at A_2 and either a $(3, 4)$ or $(4, 1)$-curve must contain a fixed point. Also, the turn at A_2 can be removed without increasing the number of fixed points. As a result, all turns can be removed and we reduce to the case $m = 0$ with $\Phi_a(X) + \Phi_b(Y) + \sigma(X, Y)$ fixed points.

Finally, to verify the formula for type T_5 we need only consider the letter $l_1 = X_b$. If $l_1 \neq b$ then the case $m = 0$ applies and since $\rho_a(\overline{X}_e) = 0$ the formula is correct. When $l_1 = b$ we use the case $m = 1$ which checks as $\Phi_b(Y) + \sigma(X, Y) = 0$.

REFERENCES

[B]   R. F. Brown, *The Lefschetz Fixed Point Theorem*, Scott-Foresman, Chicago, 1971.

[BK]   J. S. Birman, M. E. Kidwell, *Fixed points of pseudo-Anosov diffeomorphism of surfaces*, Advances in Math., **46** (1982), 217–220.

[FH]   E. Fadell and S. Husseini, *The Nielsen Number on Surfaces*, Contemporary Mathematics, Vol. 21, 59–98, AMS, Providence, 1983.

[J1]   B. J. Jiang, *Lectures on Nielsen Fixed Point Theory*, Contemporary Mathematics, Vol. 14, AMS, Providence, 1983.

[J2]   _____, *Fixed points and braids*, I, Invent. Math., **75** (1984), 69–74.

[J3]   _____, *Fixed points and braids*, II, preprint.

[J4]   _____, *Fixed Point Classes from a Different Viewpoint*, Fixed Point Theory (Sherbrooke 1980), Lecture Notes in Math., Vol. 886, 163–170, Springer-Verlag, 1981.

[J5]   _____, *Fixed points of surface homeomorphisms*, Bull. Amer. Math. Soc., **5** (1981), 176–178.

[N]   J. Nielsen, *Über die Minimilzahl der Fixpunkte bei Abbildungstypen der Ringflächen*, Math. Ann., **82** (1921), 83–93.

[W]   F. Wecken, *Fixpunktklassen*, I, II, III, Math. Ann., **117** (1941), 549–671; **118** (1942), 216–234, 544–577.

STATE UNIVERSITY OF NEW YORK
BINGHAMTON, NY 13901

AND

UNIVERSITY OF CALIFORNIA
LOS ANGELES, CA 90024

Pacific Journal of Mathematics