TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 351, Number 1, January 1999, Pages 41–62
S 0002-9947(99)01827-9

AN ALGORITHM FOR CALCULATING THE NIELSEN NUMBER ON SURFACES WITH BOUNDARY

JOYCE WAGNER

Abstract. Let $f : M \to M$ be a self-map of a hyperbolic surface with boundary. The Nielsen number, $N(f)$, depends only on the induced map $f_\#$ of the fundamental group, which can be viewed as a free group on n generators, $a_1, \dots, a_n$. We determine conditions for fixed points to be in the same fixed point class and if these conditions are enough to determine the fixed point classes, we say that $f_\#$ is W-characteristic. We define an algebraic condition on the $f_\#(a_i)$ and show that "most" maps satisfy this condition and that all maps which satisfy this condition are W-characteristic. If $f_\#$ is W-characteristic, we present an algorithm for calculating $N(f)$ and prove that the inequality $|L(f) - \chi(M)| \leq N(f) - \chi(M)$ holds, where $L(f)$ denotes the Lefschetz number of f and $\chi(M)$ the Euler characteristic of M, thus answering in part a question of Jiang and Guo.

Introduction

Given a self-map, f, on a manifold M, we wish to obtain information about the fixed point set, $\mathrm{Fix}(f) = \{x \in M | f(x) = x\}$. In particular, we are interested in finding the minimum number $MF(f)$ of fixed points among all maps homotopic to f, i.e., $MF(f) = \min\{\#\,\mathrm{Fix}(g) | g \sim f\}$.

The Nielsen number, $N(f)$, gives a lower bound for the number of fixed points of f, and because of its homotopy invariance, it is a lower bound for $MF(f)$. It is this fact that makes calculation of $N(f)$ a matter of interest.

If $N(f) = MF(f)$, we will say that f is a Wecken map. If all self-maps of M are Wecken, then M is said to be *Wecken*.

In 1941–42, Wecken published the following result:

Theorem 0.1 ([W]). *If X is a compact n-manifold where $n \geq 3$, then X is Wecken.*

In 1984–85, Jiang [J2] demonstrated that the disc with two holes, also called the pants surface, was not Wecken and was able to modify the example to all surfaces of negative Euler characteristic and obtain

Theorem 0.2 ([J3]). *A surface is Wecken if and only if its Euler characteristic is non-negative.*

Once it was known that hyperbolic surfaces are not Wecken, it became a non-trivial problem to classify those maps where $N(f) = MF[f]$. The following result was proven using Nielsen's and Thurston's work with hyperbolic surfaces.

Received by the editors December 15, 1995.
1991 *Mathematics Subject Classification.* Primary 55M20.

Theorem 0.3 ([J4], [JG])**.** *All homeomorphisms of surfaces are Wecken.*

In order to answer this classification question, we need to be able to calculate the Nielsen number. By definition, we separate the fixed points of the map f into equivalence classes and assign to each class a number, called the fixed point index. If a class has a nonzero index, we call it essential, otherwise, we say that it is inessential. The Nielsen number, $N(f)$, is then defined to be the number of essential fixed point classes. In most cases, however, calculating $N(f)$ is not easy. McCord surveys many of the existing computational methods in [Mc]. There have also been more recent results, such as the paper by Davey, Hart and Trapp [DHT], which describes an improved method for calculating the Nielsen number of maps on closed surfaces. Our goal in this paper is to develop an algorithm for calculating the Nielsen number of maps of hyperbolic surfaces with boundary for which we only need to know the induced map on the fundamental group.

In Section 1, we convert the condition of being in the same Nielsen class to a condition in the free group isomorphic to $\pi_1(M)$. We consider a map $f : C \to C$ where C is the wedge of n circles, one for each generator of the fundamental group and define $\mathbb{X}$ to be the set of images under f of these generators. In Section 2, we simplify the set $\mathbb{X}$ and determine a class for which $N(f) = |L(f)|$ where $L(f)$ is the Lefschetz number. In Section 3, we list four conditions which guarantee that two fixed points are in the same class. If those are the only conditions we need to determine the fixed point classes, we say that f is W-characteristic. We describe an easily-implemented algorithm for calculating $N(f)$ if f is W-characteristic. We define what we mean by $\mathbb{X}$ having remnant and show that most maps have remnants, in a sense that will be made precise in Theorem 3.7. Our main result is

Theorem 3.8. *If $f : C \to C$ is a map such that $\mathbb{X}$ has remnant, then f is W-characteristic.*

In Section 4, we present the algorithm and in Section 5, we look at an inequality introduced by Jiang and Guo in their paper on homeomorphisms of surfaces [JG] and show that

Theorem 5.1. *If $f : C \to C$ is W-characteristic, then $|L(f) - \chi(C)| \leq N(f) - \chi(C)$ where $\chi(C)$ is the Euler characteristic.*

In particular, if $\mathbb{X}$ has remnant, then the inequality holds. Since the Lefschetz number, the Euler characteristic and the Nielsen number are homotopy and homotopy-type invariant, we have

Corollary 5.2. *If $f : M \to M$ is a self-map of a hyperbolic surface with boundary and the induced map of the fundamental group is W-characteristic, then*

$$|L(f) - \chi(M)| \leq N(f) - \chi(M).$$

Readers who just wish to know how to calculate Nielsen numbers need only read Section 4, noting any earlier definitions as indicated. For more background information on the Nielsen number see [B], [J1], [Ki] and [N].

This paper was submitted as partial fulfillment of the requirements for the Ph.D. at the University of California at Los Angeles under the supervision of Professor Robert Brown. I wish to thank Professor Brown for his invaluable assistance and Professor Helga Schirmer for her many helpful comments.

1. When fixed points are in the same class

Suppose $f : M \to M$ is a self-map of a hyperbolic surface with boundary. Given a point x_0, the fundamental group $\pi_1(M, x_0)$ can be represented by the union C of simple closed oriented curves, $C_1, \dots, C_n$ in M, which meet only at the point x_0. Therefore, we can represent the induced map $f_\#$ of the fundamental group by a map $f_\# : C \to C$. Since M is homotopy equivalent to C and the Nielsen number is a homotopy-type invariant [J1], then $N(f) = N(f_\#)$ and so we need only consider self-maps of C. We identify $\pi_1(M, x_0)$ with the free group G on the letters a_i, $1 \le i \le n$, by associating the homotopy class of each oriented loop C_i with a_i. For simplicity of notation, we will identify any $z \in \pi_1(M, x_0)$ with its counterpart in G.

Thus, let $f : C \to C$ be any map. Because $N(f)$ is homotopy invariant, we may assume that we have chosen f so that $f(a_i) = X_i$ is a reduced word in G. We can also assume that $f(D) = x_0$ where D is the closure of a small neighborhood of x_0. The components of the complement of D are mapped onto C in the manner determined by the X_i. We can think of these complementary regions as being divided into equal intervals, where the interior of each interval is mapped homeomorphically onto $C_i - x_0$ for some i.

By construction, there is a fixed point in each interval in C_i that is sent onto itself, that is, for every a_i of a_i^{-1} that appears in X_i. The only other fixed point will be x_0. Notice that there can only be a finite number of fixed points.

We wish to determine when two fixed points x_1 on C_{i_1} and x_2 on C_{i_2} are in the same fixed point class. Let γ_j^+ and γ_j^- denote the arcs of C_{i_j} going from x_0 to x_j in the positive and negative directions, respectively. The fact that $\gamma_j^+(\gamma_j^-)^{-1} = a_{i_j}$ gives us:

Lemma 1.1. *The arcs γ_j^+ and γ_j^- satisfy the following homotopy relations:*

$$\gamma_j^- \sim a_{i_j}^{-1}\gamma_j^+ \text{ and } (\gamma_j^-)^{-1} \sim (\gamma_j^+)^{-1}a_{i_j} \text{ rel } x_0.$$

Consequently, any path γ connecting x_1 and x_2 can be written in the form

$$\gamma = (\gamma_1^+)^{-1} z \gamma_2^+$$

where $z \in \pi_1(M, x_0)$ and is therefore identified with $z \in G$. □

In the special case where $x_1 = x_0$, we write γ in the form

$$\gamma = z\gamma_2^+.$$

By definition, the fixed points x_1 and x_2 are in the same class if and only if there exists a path γ connecting the two fixed points such that $f(\gamma) \sim \gamma$ rel endpoints. By Lemma 1.1,

$$f(\gamma) = f((\gamma_1^+)^{-1}) f(z) f(\gamma_2^+).$$

To convert the relation determining fixed point classes into a condition in the free group, we would like to be able to write $f(\gamma_j^+) = W_j\gamma_j^+$ and $f(\gamma_j^-) = \overline{W}_j\gamma_j^-$ where W_j and $\overline{W}_j$ are in G. We do that as follows:

Lemma 1.2. *Suppose that x_j is the fixed point corresponding to the occurrence of $a_{i_j}^{\varepsilon_j}$ in $f(a_{i_j}) = V_j a_{i_j}^{\varepsilon_j}\overline{V}_j$ where $\varepsilon_j = \{+1, -1\}$. If we write $f(\gamma_j^+) = W_j\gamma_j^+$, then*

$$W_j = \begin{cases} V_j & \text{if } \varepsilon_j = 1, \\ V_j a_{i_j}^{-1} & \text{if } \varepsilon_j = -1. \end{cases}$$

If we write $f(\gamma_j^-) = \overline{W}_j\gamma_j^-$*, then*

$$\overline{W}_j = \begin{cases} \overline{V}_j^{-1} & \text{if } \varepsilon_j = 1, \\ \overline{V}_j^{-1} a_{i_j} & \text{if } \varepsilon_j = -1. \end{cases}$$

If $x_j = x_0$*, then* $W_j = \overline{W}_j = 1$*.*

Proof. Suppose first that $\varepsilon_j = 1$, then

$$f(\gamma_j^+)f((\gamma_j^-)^{-1}) = f(a_{i_j}) = V_j a_{i_j}\overline{V}_j = V_j\gamma_j^+(\gamma_j^-)^{-1}\overline{V}_j.$$

Since $f(x_j) = x_j$, we conclude that

$$W_j\gamma_j^+ = f(\gamma_j^+) = V_j\gamma_j^+$$

and

$$\overline{W}_j\gamma_j^- = f(\gamma_j^-) = \overline{V}_j^{-1}\gamma_j^-.$$

Therefore $W_j = V_j$ and $\overline{W}_j = \overline{V}_j^{-1}$.

Suppose now that $\varepsilon_j = -1$, then

$$f(\gamma_j^+)f((\gamma_j^-)^{-1}) = f(a_{i_j}) = V_j a_{i_j}^{-1}\overline{V}_j = V_j\gamma_j^-(\gamma_j^+)^{-1}\overline{V}_j.$$

Since $f(x_j) = x_j$, we conclude that

$$W_j\gamma_j^+ = f(\gamma_j^+) = V_j\gamma_j^-$$

and

$$\overline{W}_j\gamma_j^- = f(\gamma_j^-) = \overline{V}_j^{-1}\gamma_j^+.$$

Therefore $W_j = V_j\gamma_j^-(\gamma_j^+)^{-1} = V_j a_{i_j}^{-1}$ and $\overline{W}_j = \overline{V}_j^{-1}\gamma_j^+(\gamma_j^-)^{-1} = \overline{V}_j^{-1}a_{i_j}$.

If $x_j = x_0$, then $\gamma_j^+ = \gamma_j^- = x_0$. Since $f(x_0) = x_0$, then $f(\gamma_j^+) = \gamma_j^+$ and $f(\gamma_j^-) = \gamma_j^-$, so $W_j = \overline{W}_j = 1$. □

Lemma 1.3. *We can write* $f(a_{i_j}) = W_j a_{i_j}\overline{W}_j^{-1}$*, where* W_j *and* $\overline{W}_j$ *are defined as in Lemma* 1.2*.*

Proof. From Lemma 1.2 we have

$$\begin{aligned} f(a_{i_j}) &= f(\gamma_j^+)f((\gamma_j^{-1})^{-1}) \\ &= (W_j\gamma_j^+)((\gamma_j^-)^{-1}\overline{W}_j^{-1}) = W_j a_{i_j}\overline{W}_j^{-1}. \end{aligned}$$ □

Note that the definition of W_j and $\overline{W}_j$ depends on the exponent ε_j that is associated with the fixed point. Similarly, the fixed point index also depends on the exponent as follows:

$$i(f, x_j) = \begin{cases} -1 & \text{if } \varepsilon_j = 1, \\ +1 & \text{if } \varepsilon_j = -1 \quad \text{or} \quad x_j = x_0. \end{cases}$$

The index $i(F_i)$ of a fixed point class F_i is defined to be the sum of the indexes of all the fixed points in F_i. For the general definition of index, see [B] or [J1].

Example 1.4. If $X_1 = a_2a_1a_3^{-2}a_1^{-1}$, then there are two fixed points on the loop C_1. The first fixed point corresponds to an occurrence of a_1 and therefore $i(f, x_1) = -1$, $W_1 = a_2$ and $\overline{W}_1 = a_1a_3^2$. The second fixed point corresponds to an occurrence of a_1^{-1} and therefore $i(f, x_2) = 1, W_2 = a_2a_1a_3^{-2}a_1^{-1}$ and $\overline{W}_2 = a_1$.

We now have the following necessary and sufficient condition for two fixed points to belong to the same fixed point class.

Lemma 1.5. *Two fixed points x_1 and x_2 are in the same fixed point class if and only if there is a solution z to the equation in G:*

$$z = W_1^{-1} f(z) W_2.$$

Proof. Recall that the two fixed points are in the same fixed point class if and only if there exists a path $\gamma = (\gamma_1^+)^{-1} z \gamma_2^+$ such that $f(\gamma) \sim \gamma$ rel endpoints. From Lemma 1.1 and the definitions of the W_i, we have that

$$\begin{aligned} f(\gamma) &= f((\gamma_1^+)^{-1}) f(z) f(\gamma_2^+) \\ &= (\gamma_1^+)^{-1} W_1^{-1} f(z) W_2 (\gamma_2^+). \qquad \square \end{aligned}$$

Our goal, therefore, is to determine if there can exist such a $z \in G$, the free group on the a_i.

2. Simplifications

Given our map f, we look at the corresponding set of words $\mathbb{X} = \{X_1, \ldots, X_n\}$ and perform the following simplifications.

(S1) If there is a word $U \neq 1$ in G such that $X_i = U X_i' U^{-1}$, $X_i = U^{\pm 1}$ or $X_i = 1$ for all i and there is at least one $X_j = U X_j' U^{-1}$, then replace each X_i by $X_i' = U^{-1} X_i U$. The corresponding map f' will be homotopic to f and so $N(f) = N(f')$ since the Nielsen number is a homotopy invariant [J1].

(S2) Suppose two or more of the X_i are powers of a word U, then renumber the a_i so that $X_i = U^{r_i}$ for $m \leq i \leq n$. Let E be the bouquet of m circles and recall that C was the bouquet of n circles. We will use the commutativity property of the Nielsen number which states that if $h : E \to C$ and $g : C \to E$ are maps of finite polyhedra, then $N(g \circ h) = N(h \circ g)$ [J1]. We define maps h, g so that for $b_1, \ldots, b_m$ generating $\pi_1(E)$, we have

$$h_\#(b_i) = \begin{cases} X_i, & 1 \leq i \leq m-1, \\ U, & i = m, \end{cases}$$

$$g_\#(a_i) = \begin{cases} b_i, & 1 \leq i \leq m-1, \\ b_m^{r_i}, & m \leq i \leq n. \end{cases}$$

Since $h_\# g_\# = f_\#$ and C is a $K(\pi, 1)$, the map f is homotopic to $h \circ g$ and thus $N(f) = N(h \circ g) = N(g \circ h)$. Now, $g \circ h : E \to E$ is a map such that

$$g_\# h_\#(b_i) = \begin{cases} g_\#(X_i), & 1 \leq i \leq m-1, \\ g_\#(U), & i = m. \end{cases}$$

To simplify the notation, we will let each $g_\# h_\#(b_i) = X_i$ and $a_i = b_i$, and we are now reduced to considering $X_1, \ldots, X_m$ in the free group G on $a_1, \ldots, a_m$.

We will stop when we can no longer apply (S1) or (S2).

Theorem 2.1. *If, at some point in the simplification process, $X_i = U^{r_i}$ for all i, then $N(f) = |L(f)|$.*

Proof. We can apply (S2) with $m = 1$. Thus $g \circ h : S^1 \to S^1$ and $g_\# h_\#(b_1) = g_\#(U)$. If $\alpha_i(U)$ denotes the sum of the exponents of all the a_i^ε occurring in U, then

$$g_\#(U) = b_1^{\sum_{1 \le i \le n} r_i \alpha_i(U)}.$$

Since $g \circ h$ is a self-map of the circle having degree $\sum_{1 \le i \le n} r_i \alpha_i(U)$, then by [B] (p. 107),

$$N(g_\#) = |1 - \text{ degree } (g_\#)| = \left| 1 - \sum_{1 \le i \le n} r_i \alpha_i(U) \right|$$

which is equal to $|L(f)|$. □

Example 2.2. Let us consider the following example of Weier [We].

$$X_1 = a_1, \quad X_2 = a_4^{-1} a_2 a_4, \quad X_3 = a_3, \quad X_4 = 1.$$

We note that $X_3 = a_3$ and we may write $X_4 = a_3^0$. Therefore, we apply (S2) with $U = a_3$. Let $h_\#(b_i) = X_i$ for $1 \le i \le 3$ and

$$g_\#(a_i) = \begin{cases} b_i & \text{if } 1 \le i \le 3, \\ 1 & \text{if } i = 4. \end{cases}$$

We then look at the set $g_\# h_\#(b_i) = g_\#(X_i)$ for $1 \le i \le 3$ to obtain the new set of words

$$X_1 = a_1, \quad X_2 = a_2, \quad X_3 = a_3.$$

Since this set represents the identity map and $L(f) = -2 \ne 0$, then $N(f) = 1$.

3. Remnants and W-characteristic maps

The following is a list of sufficient conditions on the words W_j and $\overline{W}_j$ in G for two fixed points to be in the same fixed point class.

Theorem 3.1. *Two fixed points x_1 and x_2 are in the same fixed point class if one of the following occurs*:

[1.1] $W_1 = W_2$.
[1.2] $W_1 = \overline{W}_2$ *or* $\overline{W}_1 = W_2$.
[1.3] $\overline{W}_1 = \overline{W}_2$.

Proof. If we let

$$T = W_1^{-1} f(z) W_2,$$

then in each case we need to find a z such that $T = z$.

(1) Take $z = 1$, then

$$T = W_1^{-1} W_2 = 1 = z.$$

(2) The condition is symmetric, so we take $W_1 = \overline{W}_2$ and let $z = a_2^{-1}$. By Lemma 1.3,

$$T = W_1^{-1} f(a_2^{-1}) W_2 = W_1^{-1} (\overline{W}_2 a_2^{-1} W_2^{-1}) W_2 = a_2^{-1} = z.$$

(3) Let $z = a_1 a_2^{-1}$. By Lemma 1.3,

$$\begin{aligned} T &= W_1^{-1} f(a_1) f(a_2^{-1}) W_2 \\ &= W_1^{-1}(W_1 a_1 \overline{W}_1^{-1})(\overline{W}_2 a_2^{-1} W_2^{-1}) W_2 \\ &= a_1(\overline{W}_1^{-1}\overline{W}_2) a_2^{-1} = a_1 a_2^{-1} = z. \qquad \square \end{aligned}$$

Definitions 3.2. If two fixed points x_1 and x_2 satisfy [1.1], [1.2] or [1.3] of Theorem 3.1, then we say that x_1 and x_2 are *directly related*. Suppose that $x_1 = y_0, y_1, \ldots, y_m, y_{m+1} = x_2$ are fixed points. We say that $y_1, \ldots, y_m$ are *intermediate fixed points for* x_1 *and* x_2 if y_i and y_{i+1} are directly related for each $0 \leq i \leq m$. We also say that x_1 and x_2 are *related by intermediate fixed points*. Note that if x_1 and x_2 are directly related, then they are also related by intermediate fixed points with $m = 0$. If any two fixed points which are in the same fixed point class are related by intermediate fixed points, then we say that f is *W-characteristic*.

Notice that if f is W-characteristic, then a fixed point x_1 in a fixed point class F must be directly related to at least one other fixed point in F or $F = \{x_1\}$. Therefore, using W_i and $\overline{W}_i$, we can determine the fixed point classes and thus calculate $N(f)$. An example of the algorithm is presented in Section 4.

We will denote the remainder of this section to defining a certain class of maps $f : C \to C$ and proving that all such maps are W-characteristic.

Write $z = a_{j_1}^{\alpha_1} a_{j_2}^{\alpha_2} \cdots a_{j_m}^{\alpha_m}$ where each $1 \leq j_i \leq m$ and $j_i \neq j_i + 1$ for any i. Let $\Phi_{a_i}(z)$ be the number of occurrences of a_i or a_i^{-1} in z and let $A = \sum_{1 \leq i \leq m} \Phi_{a_i}(z)$. Therefore, $|z| = A$ where $|W|$ denotes the length of the reduced form of a word W.

Given z as above, we have that

$$f(z) = X_{j_1}^{\alpha_1} X_{j_2}^{\alpha_2} \cdots X_{j_m}^{\alpha_m}$$

where the right hand side of the equality is not necessarily reduced. The reduced form of $f(z)$, and thus $|f(z)|$, depend on the relationships among the X_i.

Definitions 3.3. Suppose $\mathbb{S}$ is a set of reduced words in the free group G and $T \in \mathbb{S}^{\pm 1}$ where $\mathbb{S}^{\pm 1}$ is the set of words in $\mathbb{S}$ and their inverses. Let $M(T, Z)$ be the longest initial segment of T that cancels in the product $Z^{-1}T$ for $Z \in \mathbb{S}^{\pm 1}$. Let $M(T, \mathbb{S})$ be the longest of all the $M(T, Z)$ where $Z \neq T$. Then, by definition, $M(T^{-1}, \mathbb{S})^{-1}$ will be the longest terminal segment of T that cancels in any product TZ for $Z \in \mathbb{S}^{\pm 1}$. Therefore, we can write $T = M(T, \mathbb{S})T' = T'' M(T^{-1}, \mathbb{S})^{-1}$. If we can write $T = M(T, \mathbb{S})\overline{T} M(T^{-1}, \mathbb{S})^{-1}$ where both sides of the equation are reduced and $\overline{T} \neq 1$, then we call $\overline{T}$ the *remnant of T in* $\mathbb{S}$. If every element in $\mathbb{S}$ has a remnant, then we say that $\mathbb{S}$ *has remnant*. Let

$$M(\mathbb{S}) = \{M(T, \mathbb{S})^{\varepsilon} \neq 1 | T \in \mathbb{S}^{\pm 1}, \varepsilon = \pm 1\}.$$

We will call any element in $M(\mathbb{S})$ a *maximal common factor* or MCF.

In our particular case, we are interested in the set of words $\mathbb{X} = \{X_1, \ldots, X_n\}$ in G and we say that f has remnant if $\mathbb{X}$ has remnant. We will set $P_i = M(X_i, \mathbb{X})$ and $S_i = M(X_i^{-1}, \mathbb{X})^{-1}$.

Example 3.4. Suppose $\mathbb{X} = \{X_1, X_2, X_3, X_4\}$ as follows:

$$X_1 = a_1^2 a_2^2 a_1^{-1} a_3^{-1},$$
$$X_2 = a_1^2 a_2 a_3^{-1},$$
$$X_3 = a_3 a_2^{-3} a_1^{-2},$$
$$X_4 = a_1^{-3}.$$

To determine $P_1 = M(X_1, \mathbb{X})$, we find $M(X_1, X_j^\varepsilon)$ for every $X_j^\varepsilon \neq X_1$. Since $a_1^2 a_2$ is the longest initial segment of X_1 that cancels in

$$X_2^{-1} X_1 = (a_3 a_2^{-1} a_1^{-2})(a_1^2 a_2^2 a_1^{-1} a_3^{-1}),$$

then $M(X_1, X_2) = a_1^2 a_2$. We can also see that

$$M(X_1, X_3) = 1, \quad M(X_1, X_2^{-1}) = 1,$$
$$M(X_1, X_4) = 1, \quad M(X_1, X_3^{-1}) = a_1^2 a_2^2,$$
$$M(X_1, X_1^{-1}) = 1, \quad M(X_1, X_4^{-1}) = a_1^2$$

and therefore $P_1 = a_1^2 a_2^2$.

To determine $S_1 = M(X_1^{-1}, \mathbb{X})^{-1}$ we find $M(X_1^{-1}, X_j^\varepsilon)$ for every $X_j^\varepsilon \neq X_1^{-1}$. Since there is no cancelation in the product $X_1^{-1} X_1^{-1} = (a_3 a_1 a_2^{-2} a_1^{-2})(a_3 a_1 a_2^{-2} a_1^{-2})$, then $M(X_1^{-1}, X_1) = 1$. Similarly,

$$M(X_1^{-1}, X_2) = 1, \quad M(X_1^{-1}, X_2^{-1}) = a_3,$$
$$M(X_1^{-1}, X_3) = a_3, \quad M(X_1^{-1}, X_3^{-1}) = 1,$$
$$M(X_1^{-1}, X_4) = 1, \quad M(X_1^{-1}, X_4^{-1}) = 1,$$

and so $M(X_1^{-1}, \mathbb{X}) = a_3$. Therefore, $S_1 = a_3^{-1}$ and $X_1 = [a_1^2 a_2^2] a_1^{-1} [a_3^{-1}] = P_1 a_1^{-1} S_1$. By Definition 3.3, X_1 has a remnant.

We do the same for the other X_i to get that

$$P_2 = a_1^2 a_2, \quad S_2 = a_2 a_3^{-1},$$
$$P_3 = a_3 a_2^{-1}, \quad S_3 = a_2^{-2} a_1^{-2},$$
$$P_4 = 1, \quad S_4 = a_1^{-2},$$

and therefore

$$X_2 = P_2 a_3^{-1} = a_1^2 S_2,$$
$$X_3 = P_3 S_3,$$
$$X_4 = a_1^{-1} S_4.$$

By Definition 3.3, X_4 has a remnant, but X_2 and X_3 do not, and so $\mathbb{X}$ does not have remnant.

Lemma 3.5. *If $\mathbb{X}$ has remnant and $z = a_{i_1}^{\varepsilon_1} \cdots a_{i_k}^{\varepsilon_k}$ is a non-trivial, reduced word in G, where $\varepsilon_j \in \{-1, +1\}$, then we can write $f(z) = U R_1 \cdots R_k V^{-1}$ where*

1. *$U R_1 \cdots R_k V^{-1}$ is reduced.*
2. *For all j, $R_j^{\varepsilon_j}$ is a non-trivial subword of X_{i_j}.*
3. *$U = M(X_{i_1}^{\varepsilon_1}, \mathbb{X})$ and $V = M(X_{i_k}^{-\varepsilon_k}, \mathbb{X})$.*
4. *$|f(z)| \geq |z|$.*

Proof. Since $z = a_{i_1}^{\varepsilon_1} \cdots a_{i_k}^{\varepsilon_k}$, then $f(z) = X_{i_1}^{\varepsilon_1} \cdots X_{i_k}^{\varepsilon_k}$ and therefore, by Definition 3.3, $f(z) = (P_{i_1}, \overline{X}_{i_1}, S_{i_1})^{\varepsilon_1} \cdots (P_{i_k} \overline{X}_{i_k} S_{i_k})^{\varepsilon_k}$. Since P_{i_j} and S_{i_j} are the maximal portions of X_{i_j} that can cancel in a product of the X_i's, then $\overline{X}_{i_j}^{\varepsilon_j}$ will remain uncancelled. If we let $U = M(X_{i_1}^{\varepsilon_1}, \mathbb{X})$ and $V = M(X_k^{-\varepsilon_k}, \mathbb{X})$ and R_j be the portion of $X_{i_j}^{\varepsilon_j}$ that remains uncancelled in $U^{-1}f(z)V$, then $R_j^{\varepsilon_j}$ contains $\overline{X}_{i_j}$ and so R_j is non-trivial. This implies that

$$|f(z)| \geq \sum_{1 \leq j \leq k} |R_j| \geq k = |z|. \qquad \square$$

Using Lemma 3.5, we can show that if $\mathbb{X}$ has remnant, then the set $f^k(\mathbb{X}) = \{f^k(X_1), \ldots, f^k(X_n)\}$ also has remnant for all $k \geq 1$. It is enough to prove

Lemma 3.6. *If $\mathbb{X}$ has remnant, then $f(\mathbb{X})$ has remnant.*

Proof. Suppose that $X_i \in \mathbb{X}$. Let Y and Z be any elements in $\mathbb{X}^{\pm 1}$ and write $X_i = ABC, Y = AD$ and $Z = EC$, where $A = M(X_i, Y)$ and $C = M(X_i^{-1}, Z^{-1})^{-1}$. It is possible that A and C are trivial, but since $\mathbb{X}$ has remnant, then B, D and E are non-trivial. We have that $f(X_i) = f(A)f(B)f(C), f(Y) = f(A)f(D)$ and $f(Z) = f(E)f(C)$. By Lemma 3.5(1), we can write $f(B) = UR_1 \cdots R_k V^{-1}$. Since $B_s \neq D_s$, then U is the longest initial segment of $f(B)$ that will cancel in the product $f(D)^{-1}f(B)$ and so $M(f(X_i), f(Y))$ is at most $f(A)U$. Similarly, since $B_e \neq E_e$, then $M(f(X_i)^{-1}, f(Z)^{-1})^{-1}$ is at most $V^{-1}f(C)$. Since Y and Z were arbitrary elements, then $M(f(X_i), f(\mathbb{X}))$ is at most $f(A)U$ and $M(f(X_i)^{-1}, f(\mathbb{X}))^{-1}$ is at most $V^{-1}f(C)$. Thus

$$f(X_i) = M(f(X_i), f(\mathbb{X}))\overline{f(X_i)}M(f(X_i)^{-1}, f(\mathbb{X}))^{-1}$$

where $\overline{f(X_i)}$ contains $R_1 \cdots R_k$ which is non-trivial by Lemma 3.5(2). Therefore, $f(X_i)$ has a remnant for all i, and so $f(\mathbb{X})$ has remnant. $\square$

It can also be shown that 'most' maps of hyperbolic surfaces with boundary have remnants. The following theorem and proof were suggested by Professor Robert Brown.

Theorem 3.7. *Given $\varepsilon > 0$, there exists $M > 0$ such that if $\mathbb{X} = \{X_1, \ldots, X_n\}$ is chosen at random from among all words of length $\leq M$, then the probability that $\mathbb{X}$ has remnant is greater than $1 - \varepsilon$.*

Proof. Given $\varepsilon > 0$, choose $\eta > 0$ such that $(1-\eta)^2 = 1 - \varepsilon$. Say that $\mathbb{X} > m$ if all $|X_i| > m$ for some integer m and similarly, $\mathbb{X} \leq M$ if all $|X_i| \leq M$ for some integer M. Note that

$$\begin{aligned} \text{Prob}(\mathbb{X} \text{ has remnant } |\mathbb{X} \leq M) &\geq \text{Prob}(\mathbb{X} \text{ has remnant and } \mathbb{X} > m|\mathbb{X} \leq M) \\ &= \text{Prob}(\mathbb{X} \text{ has remnant } |M < \mathbb{X} \leq \mathbb{M}) \cdot \text{Prob}(\mathbb{X} > m|\mathbb{X} \leq M). \end{aligned}$$

We first show that we can choose m so that

$$\text{Prob}(\mathbb{X} \text{ has remnant } |m < \mathbb{X} \leq M) > 1 - \eta.$$

Now

$$\begin{aligned} &\text{Prob}(\mathbb{X} \text{ has no remnant } |m < \mathbb{X} \leq M) \leq n, \\ &\text{Prob}(X_i \text{ has no remnant } |m < \mathbb{X} \leq M). \end{aligned}$$

If X_i has no remnant, then either $|P_i| \geq |X_i|/2$ or $|S_i| \geq |X_i|/2$. The probability that there exists $X_j^{\varepsilon_j} = U\overline{X}_j$ where $X_i^{\varepsilon_i} = U\overline{X}_i$, $|U| \geq |X_i|/2 \geq m/2$ and $X_j^{\varepsilon_j} \neq X_i^{\varepsilon_i}$ is less than or equal to $(2n-1)(2n)^{-m/2}$, by independence. Therefore

$$\mathrm{Prob}(X_i \text{ has no remnant } | m < \mathbb{X} \leq M) \leq 2(2n-1)(2n)^{-m/2}$$

and so

$$\mathrm{Prob}(\mathbb{X} \text{ has no remnant } | m < \mathbb{X} \leq M) \leq 2n(2n-1)(2n)^{-m/2} < (2n)^{2-m/2}.$$

We can choose m so that $(2n)^{2-m/2} < \eta$.

We will now show that given such an m we can choose M so that

$$\mathrm{Prob}(\mathbb{X} > m | \mathbb{X} \leq M) > 1 - \eta.$$

We use the fact that

$$\mathrm{Prob}(\mathbb{X} \not> m | \mathbb{X} \leq M) \leq n \, \mathrm{Prob}(|X_i| \leq m | \mathbb{X} \leq M).$$

Let $W(k)$ denote the number of reduced words in G of length k, then the number of reduced words of length less than or equal to m is

$$\sum_{0 \leq k \leq m} W(k) = 1 + \sum_{1 \leq k \leq m} 2n(2n-1)^{k-1}$$
$$= 1 + 2n(1 - (2n-1)^{m-1})/(1 - (2n-1)) = (2n-1)^{m-1}$$

and thus

$$\mathrm{Prob}(\mathbb{X} \not> m | \mathbb{X} \leq M) \leq n(2n-1)^{m-1}/(2n-1)^{M-1} = n(2n-1)^{m-M}.$$

Thus, given m, we choose M so that $n(2n-1)^{m-M} < \eta$. Therefore

$$\mathrm{Prob}(\mathbb{X} \text{ has remnant } | \mathbb{X} \leq M) > (1-\eta)^2 = 1 - \varepsilon. \qquad \square$$

Our main result in this section is to show that the conditions of Theorem 3.1 determine the fixed point classes of a given map f if $\mathbb{X}$ has remnant. In other words,

Theorem 3.8. *If $f : C \to C$ is a map such that $\mathbb{X}$ has remnant, then f is W-characteristic.*

The converse of Theorem 3.8 is not true.

Example 3.9. Suppose $\pi_1(C)$ is generated by two elements a_1, a_2 and $X_1 = a_1a_2$ and $X_2 = a_2$. Clearly $\mathbb{X} = \{X_1, X_2\}$ does not have remnant and cannot be simplified by (S1) and (S2) in Section 2. There are three fixed points and each fixed point is directly related to the other two. Therefore, there is one fixed point class and by definition, f is W-characteristic.

We also need to state that not all f are W-characteristic.

Example 3.10. Suppose $\pi_1(C)$ is generated by a_1 and a_2 and $f : C \to C$ is a map such that

$$X_1 = a_1a_2a_1a_2 \quad \text{and} \quad X_2 = a_1^{-1}.$$

There are two fixed points and x_0, and if we number them in order of occurrence, then

$$W_0 = 1, \qquad \overline{W}_0 = 1,$$
$$W_1 = 1, \qquad \overline{W}_1 = a_2^{-1}a_1^{-1}a_2^{-1},$$
$$W_2 = a_1a_2, \quad \overline{W}_2 = a_2^{-1}.$$

The only two fixed points which are directly related are x_0 and x_1. If f were W-characteristic, then x_2 would not be in the same fixed point class as x_1 and x_0. However, if we let $z = a_2$, then

$$W_1^{-1} f(z) W_2 = a_1^{-1}(a_1 a_2) = a_2 = z.$$

Therefore, x_1 and x_2 are in the same fixed point class by Lemma 1.5 and so f is not W-characteristic.

We can find a map f' which has the same Nielsen number as f and is W-characteristic. Define $\phi : C \to C$ by $\phi(a_1) = a_2^{-1} a_1$ and $\phi(a_2) = a_2$. Therefore, $\phi^{-1}(a_1) = a_2 a_1$ and $\phi^{-1}(a_2) = a_2$ and we let $f' = \phi \circ f \circ \phi^{-1}$. By commutativity of the Nielsen number, $N(f) = N(f')$. Now, $f'(a_1) = a_1 a_2$ and $f'(a_2) = a_1^{-1} a_2$ and since $\{a_1 a_2, a_1^{-1} a_2\}$ has remnant, then f' is W-characteristic.

Given any map $f : C \to C$, an interesting question would be: Under what conditions can we find a map $f' : C \to C$ such that $N(f') = N(f)$ and f' is W-characteristic?

The rest of this section will be devoted to a proof of Theorem 3.8.

Remark 3.11. If x_1 and x_2 are in the same fixed point class, then by Lemma 1.5 there exists $z \in G$ such that $z = W_1^{-1} f(z) W_2$. Since Lemma 3.5 deals with the cancelation that occurs within $f(z)$, and since W_1^{-1} and W_2 are reduced words, we need be concerned with only the cancelation occurring between $f(z)$ and each of the W_i. By Lemma 3.5, if $z = a_{i_1}^{\varepsilon_1} \cdots a_{i_k}^{\varepsilon_k}$, then $f(z) = U R_1 \cdots R_k V^{-1}$ where $U = M(X_{i_1}^{\varepsilon_1}, \mathbb{X})$ and $V = M(X_{i_k}^{-\varepsilon_k}, \mathbb{X})$. Because of the symmetry involved, we will only consider W_1^{-1}. There are three possibilities:

(R1) If W_1^{-1} does not cancel at all with $f(z)$, then $|W_1^{-1} f(z)| = |W_1^{-1}| + |f(z)|$.

(R2) Suppose W_1^{-1} cancels at least partially with U but not at all with R_1, then $|W_1^{-1} f(z)| = |W_1^{-1} U| + |R_1 \cdots R_k V|$. If (R1) or (R2) holds, we say that W_1 *cancels at most with an* MCF.

(R3) If the first two cases do not hold, then W_1 cancels at least partially with R_1. By Definition 3.3, U is the longest portion of $X_{i_1}^{\varepsilon_1}$ that cancels in any product of the form $X_j^{-1} X_{i_1}^{\varepsilon_1}$ where $X_j \neq X_{i_1}^{\varepsilon_1}$ and by Lemma 1.3, $X_m = W_1 a_m \overline{W}_1$ for some m. Therefore, since more than U is cancelled, $X_m = X_{i_1}^{\varepsilon_1}$ and the condition $z = W_1^{-1} f(z) W_2$ of Lemma 1.5 becomes

$$\begin{aligned} a_{i_1} z_1 = z = W_1^{-1} f(a_{i_1} z_1) W_2 &= W_1^{-1} X_{i_1} f(z_1) W_2 \\ &= a_{i_1} \overline{W}_1^{-1} f(z_1) W_2. \end{aligned}$$

Thus

$$z_1 = \overline{W}_1^{-1} f(z_1) W_2.$$

By the same argument, we conclude that $\overline{W}_1$ cancels at most with an MCF because if not, $f(z_1)$ must start with $X_{i_1}^{-1}$, which would imply that z_1 starts with $a_{i_1}^{-1}$. However, this would imply that $z = a_{i_1} a_{i_1}^{-1} z_2$, which is not reduced. We will refer to case (R3) by saying that W_1 *cancels with a remnant.*

Let Z_i be a general reduced word.

Lemma 3.12. *Suppose that $z = Z_1^{-1} f(z) Z_2$ where Z_1, Z_2 and $z \neq 1$ are reduced words. If Z_1 and Z_2 cancel at most with an MCF, then*

(1) *Each of Z_1 and Z_2 is either trivial or an MCF.*

(2) *The reduced form of* $Z_1^{-1}f(z)Z_2$ *consists only of powers of remnants* $\overline{X}_i$ *or, equivalently, all the MCF's are cancelled completely.*
(3) *If* $\Phi_{a_i}(z) \neq 0$, *then* $\overline{X}_i = a_i$.

Proof. We will let $A_i = \Phi_{a_i}(z)$, which counts the number of occurrences of a_i and a_1^{-1} in z. We can also suppose that $z = a_{i_1}^{\varepsilon_1} \cdots a_{i_k}^{\varepsilon_k}$ and $f(z) = UR_1 \cdots R_k V^{-1}$, as in Lemma 3.5. There are three possibilities which satisfy the hypothesis. We verify conclusions (1) and (2) in each case.

(a) If neither of the Z_i cancel at all with $f(z)$, then by (R1),

$$\sum_{1\le i\le n} A_i = |z| = |Z_1^{-1}f(z)Z_2|$$
$$= |Z_1^{-1}| + \sum_{1\le i\le n} (A_i|\overline{X}_i|) + E + |Z_2|$$

where E is the sum of the lengths of all the uncancelled portions of MCF's in $f(z)$. Since $\overline{X}_i$ is non-trivial for all i, then
(1) $|Z_i| = 0$ which implies that $Z_i = 1$ for $i = 1, 2$,
(2) $E = 0$ and so all MCF's are completely cancelled in $f(z)$.

(b) Suppose that one of the Z_i does not cancel at all and the other one cancels at least partially with an MCF. Because of the symmetry involved, we can suppose that Z_1^{-1} cancels at least partially with U and Z_2 does not cancel at all with $f(z)$. Then we have from (R2) that

$$\sum_{1\le i\le n} A_i = |Z_1^{-1}f(z)Z_2|$$
$$= |Z_1^{-1}U| + \sum_{1\le i\le k} A_i|\overline{X}_i| + (E - |U|) + |Z_2|$$

and therefore
(1) $Z_1 = U$ and $Z_2 = 1$,
(2) $E = |U|$ which implies that U was the only MCF not cancelled completely in $f(z)$ and it is cancelled completely by Z_1.

(c) If Z_1^{-1} cancels at least partially with U and Z_2 cancels at least partially with V^{-1}, then (R2) implies that

$$\sum_{1\le i\le n} A_i = |Z_1^{-1}f(z)Z_2|$$
$$= |Z_1^{-1}U| + \sum_{1\le i\le n} A_i|\overline{X}_i| + (E - |U| - |V|) + |V^{-1}Z_2|.$$

Therefore we have
(1) $Z_1 = U$ and $Z_2 = V$,
(2) $E = |U| + |V|$ which implies that all the other MCF's were cancelled completely in $f(z)$ and U and V are completely cancelled by the Z_i.

We have shown that conclusions (1) and (2) hold in all cases and it remains to verify (3). By (2), $Z_1^{-1}f(z)Z_2 = \overline{X}_{i_1}^{\varepsilon_1} \cdots \overline{X}_{i_k}^{\varepsilon_i}$ and so $|Z_1^{-1}f(z)Z_2| = \sum_{1\le i\le n} A_i|\overline{X}_i|$. Since $|z| = \sum_i A_i$, then $A_i \neq 0$ implies that $|\overline{X}_i| = 1$. Write $z = a_{i_1}^{\varepsilon_1}z_1$, then $a_{i_1}^{\varepsilon_1}z_1 = \overline{X}_{i_1}^{\varepsilon_1} \cdots \overline{X}_{i_k}$ which, since $|\overline{X}| = 1$, implies that $\overline{X}_{i_1}^{\varepsilon_1} = a_{i_1}^{\varepsilon_1}$ and thus $\overline{X}_{i_1} = a_{i_1}$. □

Proof of Theorem 3.8. If x_1 and x_2 are in the same fixed point class, then by Lemma 1.5, there exists a word z such that $z = W_1^{-1} f(z) W_2$. We will show that x_1 and x_2 are related by a set of intermediate fixed points.

If $z = 1$, then $W_1 = W_2$ [1.1], so let us now assume that z is non-trivial and distinguish cases depending on how W_1 and W_2 cancel with $f(z)$. In each case, we will find a $z' = Z_1^{-1} f(z') Z_2$ which satisfies the hypothesis of Lemma 3.12.

If W_1 and W_2 cancel at most with an MCF, then we may apply Lemma 3.12 with $Z_1 = W_1$ and $Z_2 = W_2$ and $z' = z$.

If W_1 cancels with a remnant and W_2 cancels with at most an MCF, then by (R3), $z = a_{i_1} z'$ and $z' = \overline{W}_1^{-1} f(z') W_2$ where $\overline{W}_1$ cancels at most with an MCF. We can again apply Lemma 3.12 with $Z_1 = \overline{W}_1$ and $Z_2 = W_2$.

If both W_1 and W_2 cancels with more than an MCF, then we can do the same as above to get $z = a_{i_1} z' a_{i_k}^{-1}$ where $z' = \overline{W}_1^{-1} f(z') \overline{W}_2$. We can again apply Lemma 3.12 with $Z_1 = \overline{W}_1$ and $Z_2 = \overline{W}_2$.

Therefore we write $z' = Z_1^{-1} f(z') Z_2$ where $Z_i = W_i$ or $\overline{W}_i$.

(1) If z' is trivial, then $Z_1 = Z_2$ which corresponds to one of the conditions of Theorem 3.1.

(2) Suppose now that $z' = a_{i_1}^{\varepsilon_1} \cdots a_{i_k}^{\varepsilon_k}$ is non-trivial. By Lemma 3.12(3), $X_{i_1} = P_{i_1} \overline{X}_{i_1} S_{i_1} = P_{i_1} a_{i_1} S_{i_1}$. Let x_3 be the fixed point corresponding to the 0 of a_{i_1} in X_{i_1}. If $\varepsilon_1 = 1$, that is $z' = a_{i_1} z_1$, then $f(z') = X_{i_1} f(z_1) = P_{i_1} a_{i_1} S_{i_1} f(z_1)$ and we have

$$z' = Z_1^{-1} f(z) Z_2 = Z_1^{-1} P_{i_1} a_{i_1} S_{i_1} f(z_1) Z_2.$$

Now Lemma 3.12 states that all the P_{j_i} and S_{j_i} are cancelled completely so we conclude that $Z_1 = P_{i_1} = W_3$. Similarly, if $\varepsilon_1 = -1$, then $Z_1 = S_{i_1}^{-1} = \overline{W}_3$. In either case, since $Z_1 = W_1$ or $\overline{W}_1$, then some condition of Theorem 3.1 is satisfied and x_1 and x_3 are directly related. We can write $Z_1^{-1} = X_{i_1}^{\varepsilon_1} = a_{i_1}^{\varepsilon_1} Z_3^{-1}$ where Z_3 is $S_{i_1}^{-1} = \overline{W}_3$ if $\varepsilon_1 = 1$ or $P_{i_1} = W_3$ if $\varepsilon_1 = -1$. We now have that

$$z' = a_{i_1}^{\varepsilon_1} z_1 = Z_1^{-1} X_{i_1}^{\varepsilon_1} f(z_1) Z_2 = a_{i_1}^{\varepsilon_1} Z_3^{-1} f(z_1) Z_2$$

which is equivalent to

$$z_1 = Z_3^{-1} f(z_1) Z_2.$$

Since $(z_1)_s \neq a_{i_1}^{-\varepsilon_1}$ by the reasoning at the end of (R3), then Z_3 cancels at most with an MCF. We can now repeat the entire argument with Z_3 and Z_2.

Every time we repeat (2), we get another intermediate fixed point x_i. Since z' is of finite length, there must be an m such that z_m is trivial, which implies that x_{m+2} and x_2 are directly related. This completes the proof of Theorem 3.8. □

4. The algorithm

Given a map $f : C \to C$ and the corresponding set of reduced words $\mathbb{X} = \{X_1, \ldots, X_n\}$, the algorithm consists of the following steps:

(1) Perform all possible simplification steps (S1) and (S2) and determine if the resulting map is W-characteristic.

(2) If so, determine W_j and $\overline{W}_j$ for each fixed point x_j.

(3) Determine the fixed point classes and calculate $N(f)$.

We will present an example of the calculation of $N(f)$. Suppose $\pi_1(C)$ is generated by five elements and the set $\mathbb{X} = \{X_1, \dots, X_5\}$ is given by

$$\begin{aligned}
X_1 &= a_1^2 a_2^{-2} a_5 a_4^{-1} a_3^{-1} a_1 a_3^{-1} a_2 a_1^{-1},\\
X_2 &= a_1 a_4^2 a_2^2 a_5^{-1} a_3^{-1} a_1^{-2},\\
X_3 &= a_1 a_4 a_3 a_2 a_1^{-1},\\
X_4 &= a_1^2 a_3 a_5 a_2^{-2} a_4^{-2} a_1 a_3 a_5 a_2^{-2} a_4^{-2}, a_1^{-1},\\
X_5 &= 1.
\end{aligned}$$

Step 1: *Apply* (S1) *and* (S2) *and check if the resulting map is W-characteristic*
(a) *Perform all possible* (S1) *and* (S2).

The simplifications (S1) and (S2) are defined in Section 2. Notice in our example that each $X_i = a_1 X_i' a_1^{-1}$ or $X_i = 1$, and so we can perform the simplification (S1) to get a new set

$$\begin{aligned}
X_1 &= a_1 a_2^{-2} a_5 a_4^{-1} a_3^{-1} a_1 a_3^{-1} a_2,\\
X_2 &= a_4^2 a_2^2 a_5^{-1} a_3^{-1} a_1^{-1},\\
X_3 &= a_4 a_3 a_2,\\
X_4 &= a_1 a_3 a_5 a_2^{-2} a_4^{-2} a_1 a_3 a_5 a_2^{-2} a_4^{-2},\\
X_5 &= 1.
\end{aligned}$$

Notice now that X_2, X_4 and X_5 are multiples of the subword $a_4^2 a_2^2 a_5^{-1} a_3^{-1} a_1^{-1}$. Renumber the a_i by interchanging a_2 and a_3 and thus the corresponding X_2 and X_3 to get

$$\begin{aligned}
X_1 &= a_1 a_3^{-2} a_5 a_4^{-1} a_2^{-1} a_1 a_2^{-1} a_3,\\
X_2 &= a_4 a_2 a_3,\\
X_3 &= U,\\
X_4 &= U^{-2},\\
X_5 &= U^0,
\end{aligned}$$

where $U = a_4^2 a_3^2 a_5^{-1} a_2^{-1} a_1^{-1}$. We can therefore perform the simplification (S2) with

$$g_\#(a_i) = \begin{cases} b_i & \text{if } i = 1, 2, 3,\\ b_3^{-2} & \text{if } i = 4,\\ 1 & \text{if } i = 5 \end{cases}$$

and

$$h_\#(b_i) = \begin{cases} X_i & \text{if } i = 1, 2,\\ U & \text{if } i = 3 \end{cases}$$

and look at the set

$$\begin{aligned}
g_\# h_\#(b_1) &= b_1 b_3^{-2} b_3^2 b_2^{-1} b_1 b_2^{-1} b_3 = b_1 b_2^{-1} b_1 b_2^{-1} b_3,\\
g_\# h_\#(b_2) &= b_3^{-2} b_2 b_3,\\
g_\# h_\#(b_3) &= b_3^{-4} b_3^2 b_2^{-1} b_1^{-1} = b_3^{-2} b_2^{-1} b_1^{-1}.
\end{aligned}$$

We will set $a_i = b_i$ and $X_i = g_\# h_\#(b_i)$ to get

$$X_1 = a_1 a_2^{-1} a_1 a_2^{-1} a_3,$$
$$X_2 = a_3^{-2} a_2 a_3,$$
$$X_3 = a_3^{-2} a_2^{-1} a_1^{-1}.$$

We note that we can no longer do either (S1) or (S2) and let f' be the map corresponding to $\mathbb{X}' = \{X_1, X_2, X_3\}$. If at this point, we only had one X_i, then we would apply Theorem 2.1.

(b) *Verify that f' is W-characteristic.*

We will in fact show that $\mathbb{X}'$ has remnant. The reader can check that all MCF's are contained in the brackets. See Example 3.4 for more detail on finding MCF's.

$$X_1 = [a_1] a_2^{-1} a_1 a_2^{-1} [a_3],$$
$$X_2 = [a_3^{-2}] a_2 [a_3],$$
$$X_3 = [a_3^{-2}] a_2^{-1} [a_1^{-1}].$$

Since each X_i has a portion which is not contained in any MCF, then $X = \mathbb{X}\{X_1, X_2, X_3\}$ has remnant and therefore f' is W-characteristic by Theorem 3.8.

Step 2: *Determine the W_j and $\overline{W}_j$.*

$$X_1 = \mathbf{a}_1 a_2^{-1} \mathbf{a}_1 a_2^{-1} a_3,$$
$$X_2 = a_3^{-2} \mathbf{a}_2 a_3,$$
$$X_3 = \mathbf{a}_3^{-1} \mathbf{a}_3^{-1} a_2^{-1} a_1^{-1}.$$

The letters in boldface, where a_i appears in X_i, indicate the occurrence of a fixed point. We will number these fixed points $x_1, \dots, x_5$ in order of occurrence. We therefore have six fixed points, including x_0, and we will determine the corresponding W_i and $\overline{W}_i$. See Lemma 1.2 and Example 1.4 to recall the definition of W_i and $\overline{W}_i$.

$$W_0 = 1, \qquad \overline{W}_0 = 1,$$
$$W_1 = 1, \qquad \overline{W}_1 = a_3^{-1} a_2 a_1^{-1} a_2,$$
$$W_2 = a_1 a_2^{-1}, \qquad \overline{W}_2 = a_3^{-1} a_2,$$
$$W_3 = a_3^{-2}, \qquad \overline{W}_3 = a_3^{-1},$$
$$W_4 = a_3^{-1}, \qquad \overline{W}_4 = a_1 a_2 a_3^2,$$
$$W_5 = a_3^{-2}, \qquad \overline{W}_5 = a_1 a_2 a_3.$$

Step 3: *Determine the fixed point classes and calculate $N(f') = N(f)$.*

By comparing the W_i and $\overline{W}_i$, we see that the following pairs of fixed points are directly related: $(x_0, x_1), (x_3, x_4), (x_3, x_5)$. By Theorem 3.8, we have the following fixed point classes: $F_1 = \{x_0, x_1\}, F_2 = \{x_2\}$ and $F_3 = \{x_3, x_4, x_5\}$. We calculate that $i(F_1) = 1 - 1 = 0$, $i(F_2) = -1$ and $i(F_3) = -1 + 1 + 1 = 1$ (see the comment before Example 1.4). Therefore, $N(f) = 2$.

5. The Jiang-Guo inequality

B. Jiang and G. Guo [JG] proved that the inequality $|L(f) - \chi(M)| \le N(f) - \chi(M)$ holds for surface homeomorphisms $f : M \to M$ and asked if it is true for all maps of surfaces. We will show that it is true if the induced map on the fundamental group is homotopic to a W-characteristic map and therefore, by Theorems 3.7 and 3.8, it is true for most maps.

Theorem 5.1. *If $f : C \to C$ is W-characteristic, then*

$$|L(f) - \chi(C)| \leq N(f) - \chi(C).$$

Since the Lefschetz number, the Euler characteristic and the Nielsen number are homotopy and homotopy-type invariant, we have

Corollary 5.2. *If $F : M \to M$ is a self-map of a hyperbolic surface with boundary and the induced map of the fundamental group can be viewed as a map $f : C \to C$ which is homotopic to a map which is W-characteristic, then*

$$|L(F) - \chi(M)| \leq N(F) - \chi(M).$$

For convenience, if a fixed point x_i has positive index, we will call it a *positive fixed point* and if a fixed point class has positive index, we will call it a *positive class*. Similarly, for negative index. Furthermore, we will always assume that $\varepsilon_i \in \{-1, 1\}$ and that the fixed point x_j lies on C_{i_j}. If F_i is a fixed point class, then $\mathbb{A}^+(i)$ will denote the number of positive fixed points in F_i. Let $\mathbb{A}_0^+, \mathbb{A}_p^+, \mathbb{A}_n^+$, and $\mathbb{A}^+$ be the total number of positive fixed points in the inessential fixed point classes, the positive fixed point classes, the negative fixed point classes, and all the classes, respectively. For the number of negative fixed points, we use $\mathbb{A}_0^-, \mathbb{A}_p^-, \mathbb{A}_n^-$ and $\mathbb{A}^-$. Suppose there are c^+ positive classes and c^- negative classes.

If $\pi_1(C)$ has n-generators, then $\chi(C) = 1 - n$. By definition, $L(f) = \mathbb{A}^+ - \mathbb{A}^-$ and $N(f) = c^+ + c^-$. We will look more carefully at the relationship between $\mathbb{A}^+(i)$ and $\mathbb{A}^-(i)$ for each type of class.

Since inessential fixed point classes are defined to have an index of zero, then $\mathbb{A}_0^+ = \mathbb{A}_0^-$.

In a positive class, $\mathbb{A}^+(i) > \mathbb{A}^-(i)$ and, in particular,

Lemma 5.3. *If f is W-characteristic and F_i is a positive fixed point class, then $\mathbb{A}^+(i) = \mathbb{A}^-(i) + 1$. Therefore, $\mathbb{A}_p^+ = \mathbb{A}_n^- + c^+$.*

We will postpone the proof until later.

If F_i is a negative fixed point class, then we can write $\mathbb{A}^-(i) = \mathbb{A}^+(i) + j_i$ where $j_i > 0$. If we let J be the sum of the j_i over all negative classes, then $J = \sum_i (\mathbb{A}^-(i) - \mathbb{A}^+(i)) = \mathbb{A}_n^- - \mathbb{A}_n^+$. We will prove

Lemma 5.4. *If f is W-characteristic, then $J \leq 2n + c^- - 2$.*

We will suppose Lemmas 5.3 and 5.4 hold and prove Theorem 5.1.

Proof of Theorem 5.1. The inequality $|L(f) - \chi(M)| \leq N(f) - \chi(M)$ is equivalent to showing that $N(f) \geq L(f)$ and $N(f) + L(f) - 2\chi(M) \geq 0$.

By Lemma 5.3,

$$L(f) = \mathbb{A}^+ - \mathbb{A}^- = (\mathbb{A}_0^+ - \mathbb{A}_0^-) + (\mathbb{A}_p^+ - \mathbb{A}_p^-) + (\mathbb{A}_n^+ - \mathbb{A}_n^-) = c^+ - J.$$

Since $J \geq 0$, then $L(f) \leq c^+ \leq N(f)$.

By Lemma 5.4, we have

$$\begin{aligned} N(f) + L(f) - 2\chi(C) &= (c^+ + c^-) + (c^+ - J) - 2(1 - n) \\ &= 2c^+ + (c^- + 2n - 2) - J \\ &\geq 2c^+ + (c^- + 2n - 2) - (2n + c^- - 2) = 2c^+ \geq 0. \qquad \square \end{aligned}$$

From Theorems 3.8 and 5.1 we have the following result:

Corollary 5.5. *If $f : C \to C$ is a map such that $\mathbb{X}$ has remnant, then*

$$|L(f) - \chi(C)| \leq N(f) - \chi(C).$$

The following is a corollary to Lemmas 5.3 and 5.4.

Theorem 5.6. *If f is a W-characteristic and F_i is a fixed point class of f and $i(F_i)$ is the index of F_i, then $1 - 2n \leq i(F_i) \leq 1$.*

Proof. By Lemma 5.3, we know that $i(F_i) \leq 1$. If $F_1, \ldots, F_{c^-}$ are the negative fixed point classes, then $i(F_i) = \mathbb{A}^+(i) - \mathbb{A}^-(i) = -j_i$. Since

$$J = \sum_{1 \leq i \leq c^-} j_i \leq 2n + c^- - 2$$

by Lemma 5.4, then $j_i \leq 2n + c^- - 2 - (c^- - 1) = 2n - 1$. □

We will devote the remainder of the section to proving Lemmas 5.3 and 5.4.

Lemma 5.7. *If f is W-characteristic, then no two positive fixed points can be directly related.*

Proof. If x_1 and x_2 are positive fixed points, neither of which is x_0, then for some i_1 and i_2, we can write $X_{i_1}^{\varepsilon_1} = Z_1 a_{i_1}^{-\varepsilon_1} \cdots$ and $X_{i_2}^{\varepsilon_2} = Z_2 a_{i_2}^{-\varepsilon_2} \cdots$ where

$$Z_k a_{i_k}^{-\varepsilon_k} = \begin{cases} W_k & \text{if } \varepsilon_k = 1, \\ \overline{W}_k & \text{if } \varepsilon_k = -1. \end{cases}$$

Therefore, for the conditions of Theorem 3.1 to hold, there exists ε_1 and ε_2 so that $Z_1 a_{i_1}^{-\varepsilon_1} = Z_2 a_{i_2}^{-\varepsilon_2}$. This implies that $i_1 = i_2, \varepsilon_1 = \varepsilon_2$ and $Z_1 = Z_2$ which happens only when $x_1 = x_2$.

If $x_1 = x_0$, then $W_1 = \overline{W}_1 = 1$. However, neither W_2 nor $\overline{W}_2$ can be trivial, by Lemma 1.2, and therefore x_1 and x_2 cannot be directly related. □

Define

$$T(U_i) = \{x_j \in \text{Fix}(f) | W_j = U_i \quad \text{or} \quad \overline{W}_j = U_i\}.$$

Clearly a point in $T(U_i)$ is directly related to every other point in $T(U_i)$ and therefore each $T(U_i)$ is contained in a fixed point class.

Proof of Lemma 5.3. Since F_i is positive, then $\mathbb{A}^+(i) - \mathbb{A}^-(i) \geq 1$. If $F_i = \{x_j\}$, then $i(F_i) = 1$. Otherwise, since f is W-characteristic, we can write $F_i = T(U_1) \cup \cdots \cup T(U_{m_i})$. If $T(U_j)$ did not intersect any of the other $T(U_k)$ in F_i, then none of the fixed points in $T(U_j)$ would be directly related to any of the other fixed points in F_i and therefore $F_i = U(U_j)$. We conclude, then, that there must be at least $m_i - 1$ points of F_i that are in two $T(U_j)$. By Lemma 5.7, each $T(U_i)$ contains at most one positive fixed point and so if r of the positive fixed points are contained in two of the $T(U_i)$, then $\mathbb{A}^+(i) \leq m_i - r$. Since there need to be at least $m_i - 1$ points of F_i in two $T(U_i)$, then $\mathbb{A}^-(i) \geq m_i - 1 - r$ and therefore, $\mathbb{A}^+(i) - \mathbb{A}^-(i) \leq m_i - r - (m_i - 1 - r) = 1$. □

Let $T^-(U_i)$ be the set of all the negative fixed points in $T(U_i)$ and let $\mathbb{S} = \{U_1, \ldots, U_M\}$ where $|T^-(U_i)| \geq 1$ and $T(U_i)$ is contained in a negative fixed point class. Define $P(U_i) = \{(j, \varepsilon_j) | X_j^{\varepsilon_j} = U_i a_j^{\varepsilon} \cdots\}$. In other words, each element of

$P(U_i)$ corresponds to an occurrence of U_i which is followed by a negative fixed point.

To prove Lemma 5.4, we will make a series of claims.

Claim 5.8. *If $V, Z \in \mathbb{S}$, $V \neq Z$, then $|P(V) \cap P(Z)| \leq 1$.*

Proof. If (i, ε_i) and (j, ε_j) are distinct elements of $P(V) \cap P(Z)$, then

$$X_i^{\varepsilon_i} = V a_i^{\varepsilon_i} \cdots = Z a_i^{\varepsilon_i} \cdots$$

and

$$X_j^{\varepsilon_j} = V a_j^{\varepsilon_j} \cdots = Z a_j^{\varepsilon_j} \cdots .$$

Either both V and Z are trivial or V and Z are both maximal common factors for $X_i^{\varepsilon_i}$ and $X_j^{\varepsilon_j}$. However, this implies that $V = Z$ which contradicts the hypothesis. □

Claim 5.9. *Given $P(U_1), \ldots, P(U_M)$, we can find sets $\mathbb{S}_1, \ldots, \mathbb{S}_N$ such that for each $1 \leq \alpha \leq N$, $\mathbb{S}_\alpha = P(U_{\alpha_1}) \cup \cdots \cup P(U_{\alpha_{m_\alpha}})$ and*

1. *the $\mathbb{S}_\alpha$ are disjoint,*
2. $\bigcup_{1 \leq \alpha \leq N} \mathbb{S}_\alpha = \bigcup_{1 \leq j \leq M} P(U_i)$,
3. *given any $P(U_{\alpha_k}) \subset \mathbb{S}_\alpha$, we can order $\mathbb{S}_\alpha = P(U_{\sigma(\alpha_1)}) \cup \cdots \cup P(U_{\sigma(\alpha_{m_\alpha})})$ where $\sigma(\alpha_1) = \alpha_k$ so that*

$$(*) \qquad |P(U_{\sigma(\alpha_j)}) \cap [P(U_{\sigma(\alpha_1)}) \cup \cdots \cup P(U_{\sigma(\alpha_{j-1})})]| \geq 1$$

for all $1 \leq j \leq m_\alpha$.

Proof. We will do induction on M, the number of $P(U_i)$. If $M = 1$ and we set $\mathbb{S}_1 = P(U_1)$, then $\mathbb{S}_1$ clearly satisfies (1)–(3).

Suppose that the claim is true for $M - 1$, that is, we have found sets $\mathbb{S}_1, \ldots, \mathbb{S}_N$ which contain $P(U_1), \ldots, P(U_{M-1})$ and satisfy (1)–(3).

Now consider $P(U_M)$. If $P(U_M)$ does not intersect any of the $\mathbb{S}_\alpha$, let $\mathbb{S}_{N+1} = P(U_M)$. The sets $\mathbb{S}_1, \ldots, \mathbb{S}_{N+1}$ are clearly disjoint and their union is the union of the $P(U_i)$. By the induction hypothesis, (3) is satisfied by $\mathbb{S}_1, \ldots, \mathbb{S}_N$ and it is clearly satisfied by $\mathbb{S}_{N+1}$ as well.

Suppose now that $P(U_M)$ does intersect some of the $\mathbb{S}_\alpha$. We can assume that the $\mathbb{S}_\alpha$ have been numbered so that $\mathbb{S}_{K+1}, \ldots, \mathbb{S}_N$ are the only $\mathbb{S}_\alpha$ which intersect $P(U_M)$. We define $\mathbb{T} = P(U_M) \cup \mathbb{S}_{K+1} \cup \mathbb{S}_{K+2} \cup \cdots \cup \mathbb{S}_N$ and we will show that $\mathbb{S}_1, \ldots, \mathbb{S}_K, \mathbb{T}$ satisfy (1)–(3). Since the $\mathbb{S}_1, \ldots, \mathbb{S}_N$ are disjoint and $\mathbb{S}_1, \ldots, \mathbb{S}_K$ do not intersect $P(U_M)$, then (1) is satisfied. Furthermore, since (2) holds for $\mathbb{S}_1, \ldots, \mathbb{S}_N$, it clearly holds for $\mathbb{S}_1, \ldots, \mathbb{S}_K, \mathbb{T}$ also.

By the induction hypothesis, (3) holds for $\mathbb{S}_1, \ldots, \mathbb{S}_K$ and so we only need to show that (3) holds for $\mathbb{T}$. Given any $P(U_k) \subset \mathbb{T}$ we wish to show that we can order $\mathbb{T}$ so that it starts with $P(U_k)$ and satisfies $(*)$. Suppose that $P(U_k) = P(U_M)$. Since $\mathbb{S}_\beta$ intersects $P(U_M)$ for each $K+1 \leq \beta \leq N$, there is some $P(U_i) \subset \mathbb{S}_\beta$ which intersects $P(U_M)$. Since $\mathbb{S}_\beta$ satisfies (3) by the induction hypothesis, we order $\mathbb{S}_\beta$ to start with $P(U_i)$ and to satisfy $(*)$. Now $\mathbb{T} = P(U_M) \cup \mathbb{S}_{K+1} \cup \cdots \cup \mathbb{S}_N = P(U_{\alpha_1}) \cup \cdots \cup P(U_{\alpha_m})$. Given any $P(U_{\alpha_j}) \subset \mathbb{T}$ for $j \geq 2$, we have that $P(U_{\alpha_j}) \subset \mathbb{S}_\beta$ for some β. Either $P(U_{\alpha_i})$ is the initial term in the ordering of $\mathbb{S}_\beta$ and therefore intersects $P(U_M)$, or it intersects an earlier term in the ordering of $\mathbb{S}_\beta$. In either case, $(*)$ is satisfied.

Suppose now that $P(U_k) \neq P(U_M)$. Therefore, it must be contained in some $\mathbb{S}_\gamma$ which, by (3) and the induction hypothesis, we may order to start with $P(U_k)$

and to satisfy $(*)$. We will order all the other $\mathbb{S}_\beta$ to start with an element that intersects $P(U_M)$ and to satisfy $(*)$. Now write

$$\mathbb{T} = \mathbb{S}_\gamma \cup P(U_M) \cup \bigcup_{\beta \neq \gamma} \mathbb{S}_\beta = P(U_{\alpha_1}) \cup \cdots \cup P(U_{\alpha_m}).$$

Consider any $P(U_{\alpha_i}) \subset \mathbb{T}$. Suppose $P(U_{\alpha_i}) \subset \mathbb{S}_\beta$ for some β. Either $P(U_{\alpha_i})$ is the initial term in the ordering of $\mathbb{S}_\beta$ and intersects $P(U_M)$ or it intersects an earlier term in the ordering of $\mathbb{S}_\beta$. If $P(U_{\alpha_i}) \subset \mathbb{S}_\gamma$, then it will intersect an earlier term in $\mathbb{S}_\gamma$. If $P(U_{\alpha_i}) = P(U_M)$, then $P(U_{\alpha_i})$ intersects some element in $\mathbb{S}_\gamma$ since $P(U_M)$ intersects $\mathbb{S}_\gamma$. Therefore $(*)$ holds and $\mathbb{T}$ satisfies (3). □

Claim 5.10. *If $\mathbb{S}_\alpha = P(U_{\alpha_1}) \cup \cdots \cup P(U_{\alpha_m})$ satisfies $(*)$ and U_{α_k} has the smallest length among all U_{α_i}, that is, $|U_{\alpha_k}| \leq |U_{\alpha_i}|$ for all i, then $U_{\alpha_i} = U_{\alpha_k} \cdots$.*

Proof. We will do induction on m, the number of $P(U_{\alpha_i})$ in $\mathbb{S}_\alpha$. If $m = 2$ so $\mathbb{S}_\alpha = P(U_{\alpha_1}) \cup P(U_{\alpha_2})$, then

$$X_i^{\varepsilon_i} = U_{\alpha_1} a_i^{\varepsilon_i} \cdots = U_{\alpha_2} a_i^{\varepsilon_i} \cdots$$

for some (i, ε_i). If we assume $|U_{\alpha_1}| \leq |U_{\alpha_2}|$, then $U_{\alpha_2} = U_{\alpha_1} \cdots$.

Suppose the claim is true for $m-1$. Since $P(U_{\alpha_1}) \cup \cdots \cup P(U_{\alpha_{m-1}})$ also satisfies $(*)$, then for some $k \leq m-1, U_{\alpha_i} = U_{\alpha_k} \cdots$ for all $1 \leq i \leq m-1$. Since $\mathbb{S}_\alpha$ satisfies $(*)$, then there exists at least one $j < m$ such that $|P(U_{\alpha_m}) \cap P(U_{\alpha_j})| = 1$. Therefore,

$$X_i^{\varepsilon_i} = U_{\alpha_m} a_i^{\varepsilon_i} \cdots = U_{\alpha_j} a_i^{\varepsilon_i} \cdots$$

for $(i, \varepsilon_i) = P(U_{\alpha_m}) \cap P(U_{\alpha_j})$. If $|U_{\alpha_m}| \geq |U_{\alpha_j}|$, then by the induction hypothesis, $U_{\alpha_m} = U_{\alpha_j} \cdots = U_{\alpha_k} \cdots$ and the claim holds.

If $|U_{\alpha_m}| \leq |U_{\alpha_j}|$, then $U_{\alpha_j} = U_{\alpha_m} \cdots$, but we also have $U_{\alpha_j} = U_{\alpha_k} \cdots$ by the induction hypothesis, so $U_{\alpha_m} \cdots = U_{\alpha_k} \cdots$. If $|U_{\alpha_m}| \geq |U_{\alpha_k}|$, then we are done. If not, $U_{\alpha_k} = U_{\alpha_m} \cdots$ and by the induction hypothesis, $U_{\alpha_i} = U_{\alpha_m} \cdots$ for all i. □

Claim 5.11. *If K is the total number of positive fixed points which are contained in a negative class and are contained in only one $T(U_i)$, then*

$$\sum_{1 \leq i \leq M} |P(U_i)| \geq J + M - c^- + K.$$

Proof. By definition, each element in $T^-(U_i)$ contributes exactly once to the set $P(U_i)$. Therefore $\sum_{1 \leq i \leq M} |P(U_i)| = \sum_{1 \leq i \leq M} |T^-(U_i)|$. Furthermore, since all negative fixed points contribute to either one or two of the $T^-(U_i)$, then

$$\mathbb{A}_n^- = \left| \bigcup_i T^-(U_i) \right| = \sum_i |T^-(U_i)| - \sum_{1 \leq j < k \leq M} |T^-(U_j) \cap T^-(U_k)|$$

and therefore

$$\sum_{1 \leq i \leq M} |P(U_i)| = \mathbb{A}_n^- + \sum_{1 \leq j < k \leq M} |T^-(U_j) \cap T^-(U_k)|.$$

We now wish to find a lower bound for $\sum_{1 \leq j < k \leq M} |T^-(U_j) \cap T^-(U_k)|$. As we observed in the proof of Lemma 5.3, if we write $F_i = T(U_1) \cup \cdots \cup T(U_{m_i})$, then

$\sum_{1\leq j<k\leq m_i} |T(U_j)\cap T(U_k)| \geq m_i - 1$. Suppose that there are K_i positive fixed points in F_i, each of which belongs to only one $T(U_j)$, then at most $\mathbb{A}^+(i) - K_i$ positive fixed points could account for the intersections $T(U_j)\cap T(U_k)$. Therefore, at least $m_i - 1 - (\mathbb{A}^+(i) - K_i)$ negative points must be in the intersections and so

$$\begin{aligned}\sum_{1\leq j<k\leq M} |T^-(U_j)\cap T^-(U_k)| &\geq \sum_i (m_i - 1 - (\mathbb{A}^+(i) - K_i)) \\ &= M - c^- - \mathbb{A}^+_n + K,\end{aligned}$$

where we are only summing over the negative F_i. Therefore,

$$\sum_{1\leq i\leq M} |P(U_i)| \geq \mathbb{A}^-_n + M - c^- - \mathbb{A}^+_n + K = J + M - c^- + K. \qquad \square$$

Claim 5.12. *If* $\mathbb{S}_\alpha = P(U_{\alpha_1})\cup\cdots\cup P(U_{\alpha_{m_\alpha}})$ *satisfies* $(*)$, *then*

$$|\mathbb{S}_\alpha| \geq \sum_{1\leq i\leq m} |P(U_{\alpha_i})| - (m_\alpha - 1).$$

Proof. Given $\mathbb{S}_\alpha$ ordered as above, let Γ be the undirected graph whose vertices are the $P(U_{\alpha_i})$ and which has an edge connecting $P(U_{\alpha_i})$ and $P(U_{\alpha_j})$ if and only if

$(**)$ $\quad P(U_{\alpha_i})\cap P(U_{\alpha_j}) = \varnothing$ and there is no $(k,l)\neq(i,j)$ where $k\geq i$ $l\geq j$ and $P(U_{\alpha_i})\cap P(U_{\alpha_j}) = P(U_{\alpha_k})\cap P(U_{\alpha_l})$.

We claim that Γ contains no cycles. Suppose Γ contains a cycle which has vertices $P(U_{\beta_1}),\ldots,P(U_{\beta_k})$ and edges $e_{(1,2)}, e_{(2,3)},\ldots,e_{(k-1,k)}, e_{(k,1)}$ where $e_{(i,j)}$ connects $P(U_{\beta_i})$ and $P(U_{\beta_j})$ and $k\geq 3$. Because each edge represents an intersection, $\mathbb{T} = P(U_{\beta_1})\cup\cdots\cup P(U_{\beta_k})$ satisfies $(*)$ and so by Claim 5.10, there is some r such that $U_{\beta_i} = U_{\beta_r}\cdots$ for all i. Since $P(U_{\beta_1}),\ldots,P(U_{\beta_k})$ are the vertices of a cycle, we can cyclically permute $\mathbb{T}$ and still satisfy $(*)$ and therefore we can assume that $U_{\beta_i} = U_{\beta_1}\cdots$ for all i. Suppose that $(t,\varepsilon_t) = P(U_{\beta_1})\cap P(U_{\beta_2})$. We will show that $U_{\beta_i} = U_{\beta_1}a_t^{\varepsilon_t}\cdots$ for all $2\leq i\leq k$. Since (t,ε_t) is the intersection of $P(U_{\beta_1})$ and $P(U_{\beta_2})$, then $X_t^{\varepsilon_t} = U_{\beta_1}a_t^{\varepsilon_t}\cdots = U_{\beta_2}a_t^{\varepsilon_t}\cdots$ and so $U_{\beta_2} = U_{\beta_1}a_t^{\varepsilon_t}\cdots$. Suppose that $U_{\beta_{j-1}} = U_{\beta_1}a_t^{\varepsilon_t}\cdots$. Because there exists an edge $e_{(j-1,j)}$, then $P(U_{\beta_{j-1}})$ and $P(U_{\beta_j})$ must intersect and so there is some $X_p^{\varepsilon_p} = U_{\beta_j}a_p^{\varepsilon_p}\cdots = U_{\beta_{j-1}}a_p^{\varepsilon_p}\cdots = U_{\beta_1}a_t^{\varepsilon_t}\cdots a_p^{\varepsilon_p}\cdots$. Since $|U_{\beta_1}|\leq|U_{\beta_j}|$, then $U_{\beta_j} = U_{\beta_1}a^{\varepsilon_t}a_t\cdots$. In particular, $U_{\beta_k} = U_{\beta_1}a_t^{\varepsilon_t}\cdots$. Since P_{β_1} and P_{β_k} also intersect, there is some $X_q^{\varepsilon_q} = U_{\beta_1}a_q^{\varepsilon_q}\cdots = U_{\beta_k}a_q^{\varepsilon_q}\cdots$ and so $U_{\beta_k} = U_{\beta_1}a_q^{\varepsilon_q}\cdots$. This implies that $(t,\varepsilon_t) = (q,\varepsilon_q)$ and so $P(U_{\beta_1})\cap P(U_{\beta_2}) = P(U_{\beta_1})\cap P(U_{\beta_k})$. However, this is prevented from happening by the condition $(**)$ on the edges and so Γ contains no cycles.

Therefore, Γ can be written then as a disjoint union of trees. Since Γ has m_α vertices, then it has at most $m_\alpha - 1$ edges. By Claim 5.8, each edge corresponds to a single point of intersection of the $P(U_{\alpha_i})$ and by $(**)$, there is exactly one edge for each point of intersection. Therefore there are at most $m_\alpha - 1$ distinct points of intersection and so

$$|\mathbb{S}_\alpha| \geq \sum_{1\leq i\leq m} |P(U_{\alpha_i})| - (m_\alpha - 1). \qquad \square$$

Proof of Lemma 5.4. Under a homotopy, there is a one-to-one correspondence between essential fixed point classes which preserves index [J1]. Since $-J$ is the sum of the indexes of all the negative fixed point classes, we may then apply the simplification (S1) of Section 2 without changing J, c^- or n. Therefore, we may assume that $X_i = U\overline{X}_iU^{-1}$ for all i only if $U = 1$.

Suppose we have N disjoint sets $\mathbb{S}_1, \ldots, \mathbb{S}_N$ which satisfy (1)–(3) of Claim 5.9. If $\mathbb{S}_\alpha = P(U_{\alpha_1}) \cup \cdots \cup P(U_{\alpha_{m_\alpha}})$, then by Claims 5.11 and 5.12,

$$\begin{aligned}\left|\bigcup_{1\le\alpha\le N}\mathbb{S}_\alpha\right| &= \sum_{1\le\alpha\le N}|\mathbb{S}_\alpha| \ge \sum_{1\le i\le M}|P(U_i)| - \sum_{1\le\alpha\le N}(m_\alpha - 1)\\ &\ge (J + M - c^- + K) - (M - N) = J - c^- + N + K.\end{aligned}$$

Since $\bigcup_{1\le j\le N}\mathbb{S}_j = \bigcup_{1\le i\le M}P(U_i)$ is the set of pairs of the form (i, ε_i) where $1 \le i \le n$ and $\varepsilon_i = 1$ or -1, then $|\bigcup_j \mathbb{S}_j| \le 2n$. Therefore, $2n \ge J - c^- + N + K$ and so

$$J \le 2n + c^- - (N + K) \le 2n + c^- - 1.$$

We will show that we cannot get equality.

If $J = 2n + c^- - 1$, then

$$\begin{aligned}2n \ge \left|\bigcup_{1\le j\le N}\mathbb{S}_j\right| &\ge J - c^- + N + K\\ &= (2n + c^- - 1) - c^- + N + K = 2n + N + K - 1\end{aligned}$$

where $N \ge 1$ and $K \ge 0$. Therefore, $N = 1, K = 0$ and $|\bigcup_i P(U_i)| = 2n$. In other words, all the $P(U_i)$ combine into one set $\mathbb{S}_1$ and so there is only one negative fixed point class F. Claim 5.10 implies that if U_k is the shortest of all the U_i, then $U_i = U_k \cdots$. Since $|\bigcup_i P(U_i)| = 2n$, then for every $1 \le r \le n$, there is some i and j such that $(r, 1) \in P(U_i)$ and $(r, -1) \in P(U_j)$. Therefore $X_r = U_i \cdots U_j^{-1} = U_k \cdots U_k^{-1}$ for every $1 \le r \le n$. By the assumption at the beginning of this proof, $U_k = 1$ and so $T(1) \subset F$. However, $x_0 \in T(1) \subset F$ is a positive fixed point which is only contained in $T(1)$ and so by definition, $K \ge 1$, a contradiction. □

References

[B] R. Brown, *The Lefschetz Fixed Point Theorem*, Scott-Foresman, Glenview, IL, 1971. MR **44:**1023

[DHT] O. Davey, E. Hart and K. Trapp, *Computation of Nielsen numbers for maps of closed surfaces*, Trans. Amer. Math. Soc. **348** (1996), 3245–3266. MR **97g:**55001

[J1] B. J. Jiang, *Lectures on Nielsen fixed point theory*, Contemporary Mathematics **14** (1983). MR **84f:**55002

[J2] ______, *Fixed points and braids*, I, Invent. Math. **75** (1984), 69–74. MR **85h:**55003

[J3] ______, *Fixed points and braids*, II, Math. Ann. **272** (1985), 249–256. MR **86i:**55004

[J4] ______, *Fixed points of surface homeomorphisms*, Bull. Amer. Math. Soc. **5** (1981), 176–178. MR **82m:**55005

[JG] B. Jiang and J. Guo, *Fixed points of surface diffeomorphisms*, Pacific J. Math. **160** (1993), 67–89. MR **94d:**57037

[Ki] T. Kiang, *The Theory of Fixed Point Classes*, Springer-Verlag, Berlin-New York, 1989. MR **90h:**55002

[Mc] C. McCord, *Computing Nielsen Numbers*, Nielsen Theory and Dynamical Systems, Contemporary Mathematics **152** (1993), 249–267. MR **94e:**55004

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

[W] F. Wecken, *Fixpunktklassen*, III, Math. Ann. **118** (1942), 544–577. MR **5:**275b

[We] J. Weier, *Über Probleme aus der Topologie der Ebene und der Flächen*, Math. Japon. **4** (1956), 101–105. MR **20:**4269

Department of Mathematics, California State University, Long Beach, California 90840

E-mail address: pslavich@aol.com