REPRESENTING HOMOLOGY AUTOMORPHISMS OF NONORIENTABLE SURFACES

by

John D. McCarthy and Ulrich Pinkall

Max-Planck-Institut
für Mathematik
Gottfried-Claren-Str. 26
D-5300 Bonn 3

Sonderforschungsbereich
Theoretische Mathematik
Beringstr. 4
D-5300 Bonn 1

MPI/SFB 85-11

Representing Homology Automorphisms of Nonorientable Surfaces

John D. McCarthy and Ulrich Pinkall

Introduction

It is well known that the diffeomorphisms on a closed, connected, orientable surface of genus g, M_g, induce the full group of automorphisms of $H_1(M_g,\mathbb{Z})$ which preserve the associated intersection pairing. With respect to the standard basis of $H_1(M_g,\mathbb{Z})$, this group is identified with the group of integer symplectic matrices, $Sp(2g,\mathbb{Z})$. Clebsch and Gordon discovered generators for $Sp(2g,\mathbb{Z})$ in 1866. Consequently, in 1890 Burkhardt [Bu] gave the first proof of this fact by showing that these generators are induced by diffeomorphisms of M_g. A similar algebraic proof involves the set of four generators discovered by Hua and Reiner [HR], [Bi]. Meeks and Patrusky [MP] gave a topological proof in 1978.

In the case of a closed, connected, nonorientable surface of genus p, F_p, there is only a $\mathbb{Z}_2$-valued intersection pairing. (Here, the genus of a nonorientable surface is defined to be the number of projective planes in a connected sum decomposition.) Nevertheless, we shall show in this article that the above result extends in a natural way to nonorientable surfaces. More precisely, we shall prove the following theorem.

Theorem 1 If L is an automorphism of $H_1(F_p,\mathbb{Z})$ which preserves the associated $\mathbb{Z}_2$-valued intersection pairing, then L is induced by a diffeomorphism of F_p .

Our arguments are essentially algebraic and elementary. After describing the action of certain diffeomorphisms on $H_1(F_p,\mathbb{Z}_2)$, we prove the following theorem.

Theorem 2 If L is an endomorphism of $H_1(F_p,\mathbb{Z}_2)$ which preserves the associated $\mathbb{Z}_2$- valued intersection pairing, then L is induced by a diffeomorphism which is a product of Dehn twists.

We then compute the action of certain crosscap slides (the Y-homeomorphisms of Lickorish [L], [Ch]) on $H_1(F_p,\mathbb{Z})$. By a purely algebraic argument, similar to the standard argument for establishing the generation of $GL(n,\mathbb{Z})$ by elementary matrices and permutation matrices, we deduce the following theorem.

Theorem 3 If L is an automorphism of $H_1(F_p,\mathbb{Z})$ which induces the trivial automorphism of $H_1(F_p,\mathbb{Z}_2)$, then L is induced by a diffeomorphism which is a product of crosscap slides.

Actually, we shall prove more precise versions of all three of the above theorems, versions which provide finite sets of generators (Theorem 3.1, Theorem 2.2 and Theorem 1.1). In particular, in the case of Theorem 1, we shall prove that L is induced by a diffeomorphism of F_p which is generated by a specific set of four maps, a crosscap transposition, a crosscap p-cycle, a Dehn twist and a crosscap slide (Theorem 3.1).

Here is an outline of the paper. In section 1, we describe the collection of diffeomorphisms used in the proof of Theorem 2, compute their action on $H_1(F_p,\mathbb{Z}_2)$ and prove Theorem 2. In section 2, we describe the corresponding collection of crosscap slides used in the proof of Theorem 3, compute their action on $H_1(F_p,\mathbb{Z})$ and deduce Theorem 3. Finally, in section 3, we deduce Theorem 1 from Theorems 2 and 3.

Section 1

Consider the following model for F_p. Let S_p be a sphere with p open discs removed. We denote the boundary components of S_p by $b_1,\ldots,b_p$. To each component, b_i, we attach a Moebius band, M_i, with core circle, c_i. The resulting surface, F_p, is depicted in figure 1 (for p=3). The Moebius bands are drawn as "crosscaps".

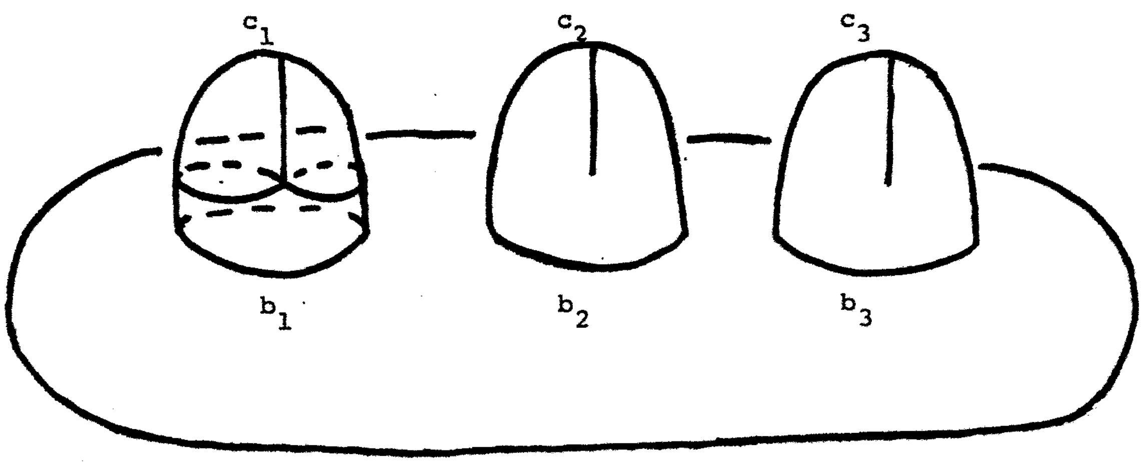

Figure 1

If $p\geq 2$, choose a diffeomorphism of F_p, τ, which exchanges the pairs (M_1,c_1) and (M_2,c_2) and fixes each of the remaining crosscaps. We refer to τ as a crosscap transposition. The action of τ is depicted in figure 2.

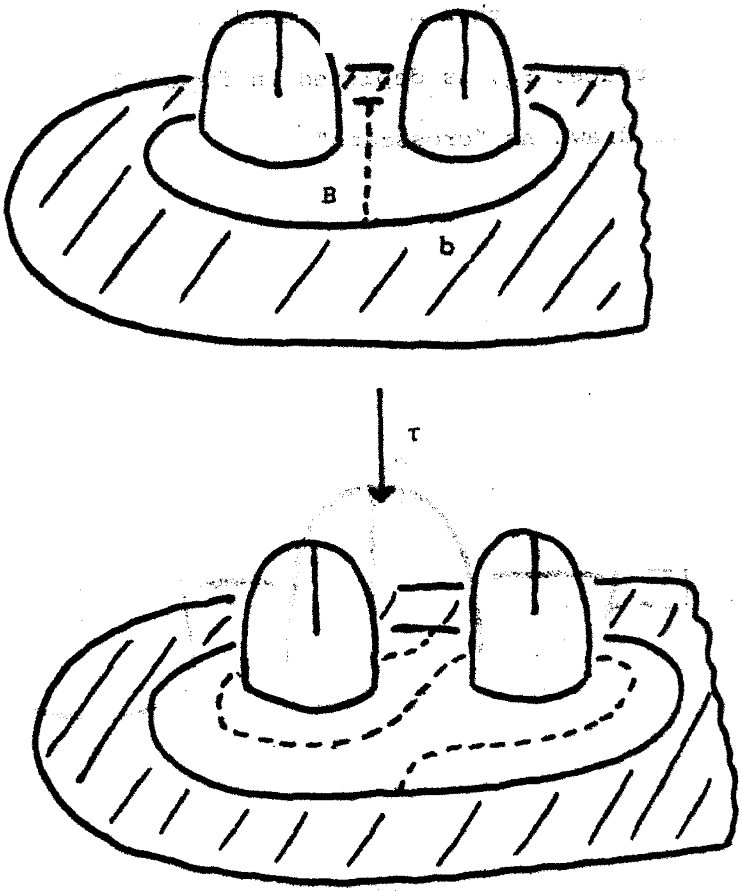

Figure 2

In a similar manner, if $p \geq 3$, choose a diffeomorphism of F_p, Ψ, which permutes the crosscaps cyclically in the given order. We refer to Ψ as a crosscap p-cycle. (We could, of course, define Ψ as an appropriate product of crosscap transpositions.) The action of Ψ is depicted in figure 3 (for $p=3$).

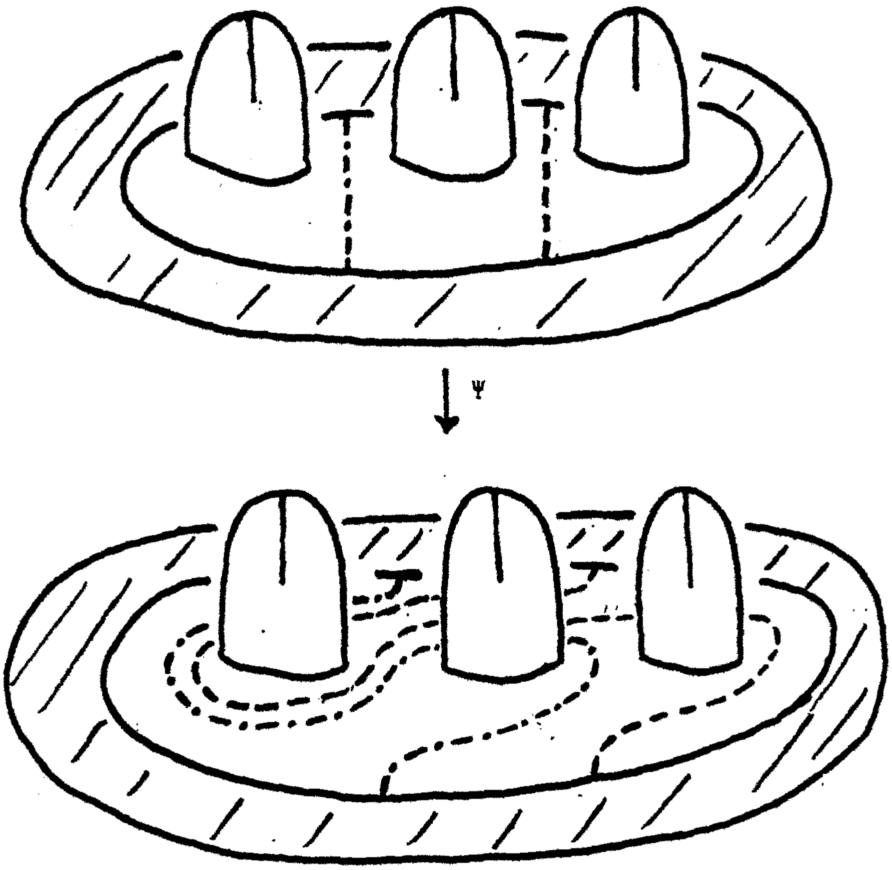

Figure 3

Finally, if $p \geq 4$, we construct a simple closed curve, d, which "runs ce around each Moebius band, M_1, M_2, M_3, and M_4, in the given order", as figure 4.

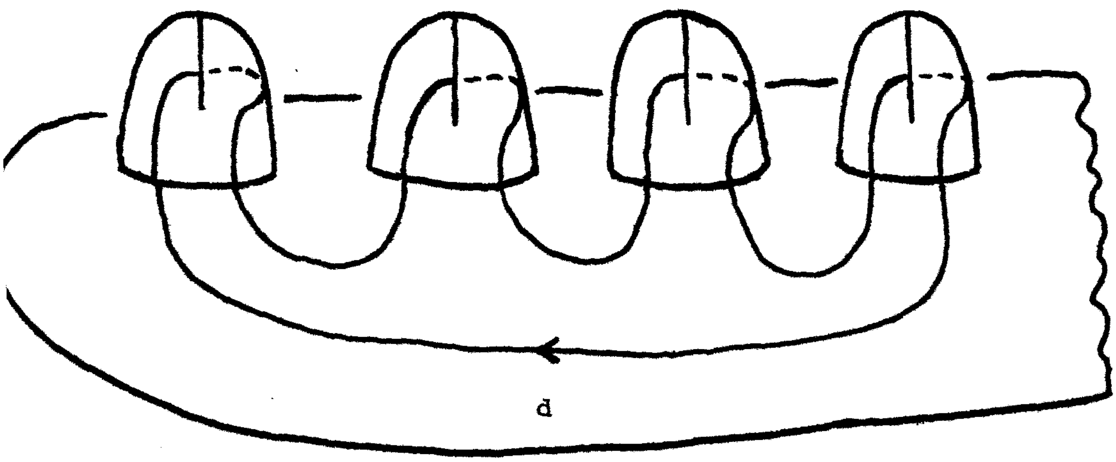

Figure 4

Since d is orientation preserving and, hence, two-sided, we may construct the Dehn twist about d ([L]), δ.

Remark For nonorientable surfaces, it is not possible to distinguish between right and left twists. Nevertheless, we could choose an orientation of a regular neighborhood of d and let δ denote the right twist about d with respect to this orientation.

The $\mathbb{Z}_2$-homology classes represented by the cores, which we also denote as $c_1,\ldots,c_p$, form a $\mathbb{Z}_2$ basis for $H_1(F_p,\mathbb{Z}_2)$.

(1.1) $H_1(F_p,\mathbb{Z}_2) = \langle c_1,\ldots,c_p\rangle$.

The $\mathbb{Z}_2$-valued intersection pairing, $\langle\ ,\ \rangle$, is given by the following conditions.

(1.2) $\langle c_i,c_j\rangle = 0 \qquad 1\le i\le p \quad 1\le j\le p \quad i\ne j$

$\langle c_i,c_i\rangle = 1$.

(In other words, with respect to the given basis, the pairing is identified with the standard inner product on $\mathbb{Z}_2^p$.)

The characteristic class, c, is given as the sum of the basis.

(1.3) $c = c_1 +\ldots+ c_p$.

It is easy to check that it is the unique $\mathbb{Z}_2$-class which satisfies the following identity.

(1.4) $\langle c,x\rangle = \langle x,x\rangle \qquad x \in H_1(F_p,\mathbb{Z}_2)$.

The $\mathbb{Z}_2$-homology class of d, which we also denote as d, is given by the following sum.

(1.5) $d = c_1 + c_2 + c_3 + c_4$.

The actions of τ, Ψ, and δ on $H_1(F_p,\mathbb{Z}_2)$ are given by the following conditions.

(1.6) $\tau_*(c_1) = c_2$ $\quad 3 \le i \le p$

$\tau_*(c_2) = c_1$

$\tau_*(c_i) = c_i$.

(1.7) $\Psi_*(c_i) = c_{i+1}$ $\quad 1 \le i \le p$ (modulo p).

(1.8) $\delta_*(x) = x + \langle d,x \rangle d$ $\quad x \in H_1(F_p,\mathbb{Z}_2)$.

Remark The $\mathbb{Z}$-homology classes corresponding to the cores, which we also denote as $c_1,\ldots,c_p$, generate $H_1(F_p,\mathbb{Z})$. Formulas (1.6) and (1.7) also hold for the actions of τ and Ψ respectively on $H_1(F_p,\mathbb{Z})$. A formula similar to (1.8) holds for the corresponding action of δ. We must replace $\langle d,x \rangle$ with a "local intersection number", $n_d(x)$, which is defined with respect to a given local orientation in a neighborhood of d.

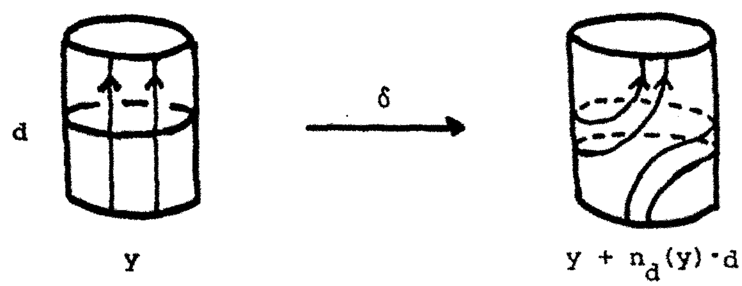

Figure 5

Suppose that L is an endomorphism of $H_1(F_p,\mathbb{Z}_2)$ which preserves the pairing. Since the pairing is nondegenerate, L must be an automorphism of $H_1(F_p,\mathbb{Z}_2)$.

Let k be the least positive integer which satisfies the following condition.

(1.9) $L(c_j) = c_j \qquad k<j\leq p$.

Observe that k=p if L fixes no basis vector and k<p otherwise.

(1.10) $1\leq k\leq p$.

It follows from the assumption that L preserves the pairing, (1.2), (1.3) and (1.9) that $L(c_k)$ has the following form.

(1.11) $L(c_k) = c_{i_1} + \ldots + c_{i_\ell} \qquad 1\leq i_1,\ldots,i_\ell\leq k$.

Again, since $\langle L(c_k), L(c_k)\rangle = \langle c_k,c_k\rangle = 1$, we conclude that ℓ is odd.

(1.12) ℓ is an odd integer.

Since the coefficients are in $\mathbb{Z}_2$, we obtain a unique expression (1.11) by imposing the following condition.

(1.13) $i_j < i_{j+1} \qquad 1\leq j<k$.

In particular, ℓ is bounded by k.

(1.14) $1\leq\ell\leq k$.

From (1.6) and (1.7), we observe that τ_* and Ψ_* generate a group of automorphisms of $H_1(F_p,\mathbb{Z}_2)$ which acts by permutations of the given basis, $\{c_1,\ldots,c_p\}$. In fact, they generate the full group of such automorphisms, which is isomorphic to the symmetric group on p-symbols, Σ_p. (Note that under the above conventions this assertion holds for all values of p.) Hence, we identify this group with Σ_p. Likewise, for each integer k with $1\leq k<p$, we identify the subgroup of Σ_p which fixes $c_{k+1},\ldots,c_p$ with Σ_k.

We are now able to state and prove a more precise version of Theorem 2. (For the sake of uniformity, we adopt the conventions that τ is the identity map when $p\leq 1$, Ψ is the identity map when $p\leq 2$, and δ is the identity map when $p\leq 3$.)

<u>Theorem 1.1</u> If L is an endomorphism of $H_1(F_p,\mathbb{Z}_2)$ which preserves the associated $\mathbb{Z}_2$-valued intersection pairing, then L is generated by τ_*, Ψ_* and δ_*.

<u>Proof:</u> Let L be an endomorphism of $H_1(F_p,\mathbb{Z}_2)$ as above. The proof is by induction on the lexicographical ordering on the pair (k,ℓ) or $(k(L),\ell(L))$.

(1) $(k,\ell) < (k',\ell')$ if and only if either (i) $k<k'$
or (ii) $k=k'$ and $\ell<\ell'$.

(Due to the inequality (1.14), this is ordinary induction on the integer $[(k-1)k+2\ell]/2$.)

The initial step of the induction is given by $k=\ell=1$, in which case L is the identity and the theorem is immediate. Henceforth, we assume that $k \geq 2$.

Suppose that $\ell = k$. By (1.11) and (1.13), it follows that $L(c_k) = c_1 + \ldots + c_k$. Applying (1.9), we conclude that $L(c_k + \sum_{k<j\leq p} c_j) = c$. From the characteristic property of c (1.4) and the assumption that L preserves the pairing, it follows that L preserves c. Since L is invertible, we conclude that $c_k + \sum_{k<j\leq p} c_j = c$ or $k=1$. This contradicts our assumption on k. Hence, $1 \leq \ell < k$.

Suppose that $\ell = 1$. Choose a permutation in Σ_k, P, such that $PL(c_k) = c_k$. Since P is in Σ_k, $PL(c_j) = c_j$ for each $j > k-1$. By the induction hypothesis, PL is generated by τ_*, Ψ_* and δ_*. By the previous remarks about Σ_k, L is generated by τ_*, Ψ_* and δ_*.

Suppose that $\ell = 3$. Choose a permutation in Σ_k, Q, such that $QL(c_k) = c_2 + c_3 + c_4$. By (1.2) and (1.8), it follows that $\delta_* QL(c_k) = c_1$. Since $1 \leq \ell < k$, we observe that $k \geq 4$. Therefore, by the same formulas and the fact that Q is in Σ_k, we compute that $\delta_* QL(c_j) = c_j$ for each $j > k$. By the induction hypothesis, $\delta_* QL$ (and hence L) is generated by τ_*, Ψ_* and δ_*.

Finally, suppose that $\ell \geq 5$. Choose a permutation in Σ_k, R, such that $RL(c_k) = c_2 + c_3 + c_4 + c_5 + \ldots + c_{\ell+1}$. As before, we conclude that $\delta_* RL$ (and hence L) is generated by τ_*, Ψ_* and δ_*.

This completes the proof of Theorem 1.1.

With respect to the basis given above, we can identify $H_1(F_p, \mathbb{Z}_2)$ with the standard vector space over $\mathbb{Z}_2$ of dimension p, $(\mathbb{Z}_2)^p$. In addition, as previously observed, under this identification, the intersection pairing on $H_1(F_p, \mathbb{Z}_2)$ is the standard inner product on $(\mathbb{Z}_2)^p$. Hence, Theorem 1.1 is a statement about the generation of $O(p, \mathbb{Z}_2)$, the group of orthogonal transformations of $(\mathbb{Z}_2)^p$.

The "transposition", τ_*, is also a transvection.

(1.15) $\tau_*(x) = x + \langle c_1+c_2, x\rangle (c_1+c_2) \qquad x \in H_1(F_p, \mathbb{Z}_2).$

By our remarks concerning Σ_p, Ψ_* is a product of transpositions. In fact, it is generated by the following transvections.

(1.16) $T_i(x) = x + \langle c_i+c_{i+1}, x\rangle (c_i+c_{i+1}) \qquad x \in H_1(F_p, \mathbb{Z}_2) \qquad 1 \le i \le p-1.$

Finally, we have already observed that δ_* is a transvection (1.8), which we shall denote as D.

(1.17) $D(x) = x + \langle d, x\rangle d \qquad x \in H_1(F_p, \mathbb{Z}_2).$

Hence, we have the following immediate corollary.

Corollary 1.2 $O(p, \mathbb{Z}_2)$ is generated by the p transvections, $T_1, \ldots, T_{p-1}$, and D.

Remark All of the above transvections are induced by Dehn twists. For example, τ_* is induced by a Dehn twist about any simple closed curve which represents c_1+c_2. Such a curve can be constructed in the same manner as d. In fact, this construction can be used to show that every invertible transvection of $H_1(F_p, \mathbb{Z}_2)$ is induced by a Dehn twist. (The reader can easily check that a transvection in the direction of a given $\mathbb{Z}_2$-homology class is invertible if and only if the given class is orientation preserving.)

Hence, the corollary above is very similar to, although considerably simpler than, the corresponding statement for Sp(2g,$\mathbb{Z}$). Furthermore, the original statement of Theorem 2 which was formulated in the introduction follows immediately.

Section 2

In this section, we consider the same model for F_p which was used in section 1 (figure 1). However, we need to introduce some additional notation.

If $p \geq 2$, choose a curve, b, which separates M_1 and M_2 from the other crosscaps. Let B be the component of $F_p \setminus b$ which contains M_1 and M_2, as in figure 2.

There is a standard diffeomorphism of F_p, σ, which is supported on B and known as a Y-homeomorphism ([L]). This diffeomorphism is constructed by "sliding the crosscap, M_1, over the crosscap, M_2, and back to itself." Hence, we shall refer to σ as a crosscap slide.

We may describe σ more explicitly as follows. Consider $B \setminus M_1$, which is a Moebius band with one open disc removed as in figure 6.

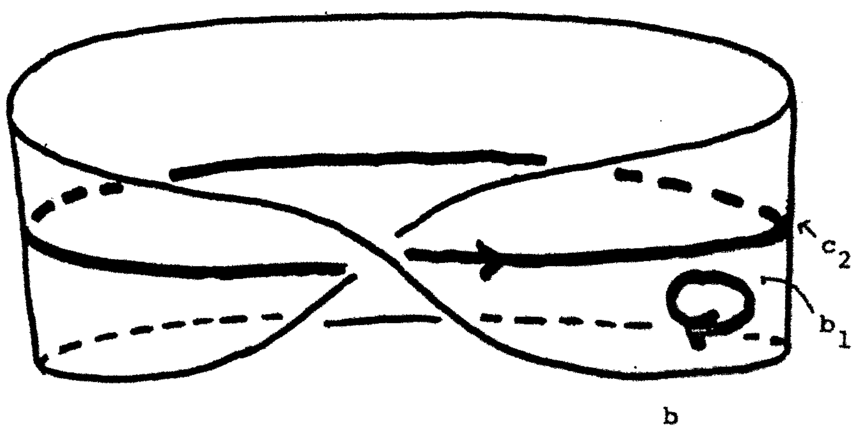

Figure 6

Construct a diffeomorphism of $B\setminus M_1$, which fixes b, by "dragging b_1 once around c_2". Of course, b_1 comes back to itself with a change of orientation as in figure 7.

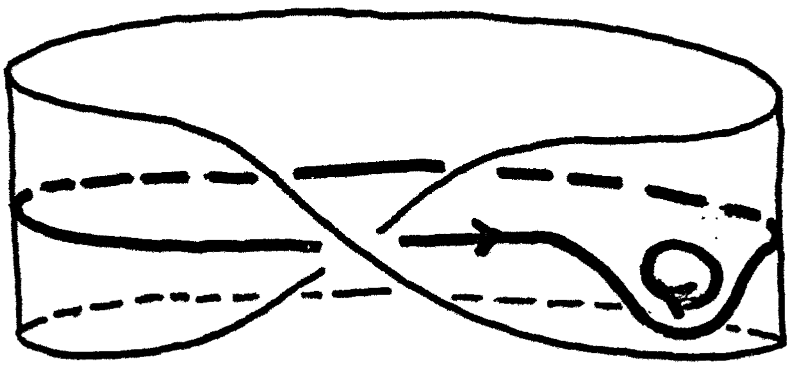

Figure 7

Hence, we may extend this map to a diffeomorphism of F_p which is supported on B, preserves M_1 and reverses the orientation of the core circle, c_1. This is the crosscap slide, σ.

From this description and figure 7, we easily calculate the action of σ on $H_1(F_p,\mathbb{Z})$.

(2.1) $\sigma_*(c_1)=-c_1$ $\quad\quad 2<j\leq p$

$\sigma_*(c_2)=c_2+b_1=2c_1+c_2$

$\sigma_*(c_j)=c_j$.

We previously observed that the actions of τ and Ψ on $H_1(F_p,\mathbb{Z})$ are given by (1.6) and (1.7) of Section 1. Hence, as for $H_1(F_p,\mathbb{Z}_2)$, τ_* and Ψ_* generate a group which we identify with Σ_p. By conjugating σ_* by elements of Σ_p, we obtain the following collection of automorphisms of $H_1(F_p,\mathbb{Z})$ which are induced by appropriate crosscap slides.

(2.2) $\quad Y_{ij}(c_i)=-c_i \qquad 1\le i\le p-1 \quad 1\le j\le p \quad 1\le k\le p$

$Y_{ij}(c_j)=2c_i+c_j \qquad i\ne j \quad j\ne k \quad k\ne i$

$Y_{ij}(c_k)=c_k.$

Note Each of these automorphisms acts trivially on $H_1(F_p,\mathbb{Z}_2)$.

The characteristic class of $H_1(F_p,\mathbb{Z})$, c, is given as the sum of the generators.

(2.3) $\quad c=c_1+\ldots+c_p.$

It is the unique class of order 2. In particular, it is preserved by every automorphism of $H_1(F_p,\mathbb{Z})$.

Let R_p denote the quotient, $H_1(F_p,\mathbb{Z})/\mathbb{Z}_2$, where $\mathbb{Z}_2$ is the span of c. As a $\mathbb{Z}$-module, $H_1(F_p,\mathbb{Z})$ has the following presentation.

(2.4) $\quad H_1(F_p,\mathbb{Z})=\langle c_1,\ldots,c_p \mid 2(c_1+\ldots+c_p)=0\rangle .$

It is immediate that R_p is a free $\mathbb{Z}$-module with basis given by the images of $c_1,\ldots,c_{p-1}$ in R_p. (We shall denote the images of $c_1,\ldots,c_p$ in R_p by the same symbols.)

(2.5) $\quad R_p=\langle c_1,\ldots,c_p \mid c_1+\ldots+c_p=0\rangle$

$R_p=\langle c_1,\ldots,c_{p-1}\rangle .$

Since every automorphism of $H_1(F_p,\mathbb{Z})$ preserves c, R_p is a characteristic quotient of $H_1(F_p,\mathbb{Z})$. That is, every automorphism of $H_1(F_p,\mathbb{Z})$, L, covers an automorphism of R_p, $\bar{L}$.

$$(2.6)\qquad \begin{array}{ccc} H_1(F_p,\mathbb{Z}) & \xrightarrow{\ L\ } & H_1(F_p,\mathbb{Z}) \\ \downarrow & & \downarrow \\ R_p & \xrightarrow{\ \bar{L}\ } & R_p \end{array}$$

In particular, the automorphisms defined above in (2.2) induce the following automorphisms of R_p.

(2.7) $\overline{Y_{ij}}(c_i)=-c_i$ $\qquad$ $1\le i\le p-1$ $\quad$ $1\le j\le p-1$ $\quad$ $1\le k\le p-1$

$\overline{Y_{ij}}(c_j)=2c_i+c_j$ $\qquad$ $i\ne j$ $\quad$ $j\ne k$ $\quad$ $k\ne i$

$\overline{Y_{ij}}(c_k)=c_k$.

(2.8) $\overline{Y_{ip}}(c_i)=-c_i$ $\qquad$ $1\le i\le p-1$ $\quad$ $1\le k\le p-1$ $\quad$ $i\ne k$

$\overline{Y_{ip}}(c_k)= c_k$.

We refer to the automorphisms in (2.8) as <u>sign change transformations of R_p</u>.

(2.9) $F_i=\overline{Y_{ip}}$ $\qquad$ $1\le i\le p-1$.

From these automorphisms of R_p, we generate the following <u>elementary transformations of R_p</u>.

(2.10) $E_{ij}=\overline{Y_{ij}}\circ\overline{Y_{ip}}$ $\qquad$ $1\le i\le p-1$ $\quad$ $1\le j\le p-1$ $\quad$ $1\le k\le p-1$

$E_{ij}(c_i)=c_i$ $\qquad$ $i\ne j$ $\quad$ $j\ne k$ $\quad$ $k\ne i$

$E_{ij}(c_j)=2c_i+c_j$

$E_{ij}(c_k)=c_k$.

The reader will recognize these transformations as the squares of the standard elementary transformations of $\mathbb{Z}^{p-1}$ (under the obvious identification of R_p with $\mathbb{Z}^{p-1}$), the so-called row and column operations. In the terminology of matrices, these standard elementary transformations together with the permutation matrices generate $GL(n,\mathbb{Z})$. A standard proof of this fact uses a type of Euclidean algorithm on the rows and columns of the corresponding matrices. Our proof of Theorem 2 uses the same idea. In order to employ the appropriate Euclidean algorithm, we shall need the following elementary lemma.

Lemma 2.1 Let a and b be nonzero integers. Suppose that a is odd and b is even. Then one of the following inequalities must be satisfied:

(i) $|a-2b|<|a|$

(ii) $|a+2b|<|a|$

(iii) $|b-2a|<|b|$

(iv) $|b+2a|<|b|$.

Proof: Suppose a and b are as above and none of the inequalities is satisfied. In other words:

(1) $|a|\leq|a+2\varepsilon b|$, $|b|\leq |b+2\varepsilon a|$, $\varepsilon=\pm 1$.

From this it follows that:

(2) $|a|^2\leq |a|^2-4|a||b|+4|b|^2$,

$|b|^2\leq |b|^2-4|a||b|+4|a|^2$.

This latter pair of inequalities implies:

(3) $|a||b|\leq|b|^2$, $|a||b|\leq|a|^2$.

Since a and b are nonzero:

(4) $|a|=|b|$.

This is clearly contradictory to the hypotheses. Hence, the lemma is established.

We are now prepared to state and prove a more precise form of Theorem 3.

Theorem 2.2 Let L be an automorphism of $H_1(F_p,\mathbb{Z})$ which acts trivially on $H_1(F_p,\mathbb{Z}_2)$. Then L is generated by $\{Y_{ij} \mid 1\leq i\leq p-1 , 1\leq j\leq p, i\neq j\}$.

Proof: Suppose that L is as above and the induced automorphism of R_p, $\bar{L}$, is the identity.

(1) $L(c_i)\equiv c_i$ (modulo c) $1\leq i\leq p$.

For any index, i, either $L(c_i)=c_i$ or $L(c_i)=c_i+c$. Since L acts trivially on $H_1(F_p,\mathbb{Z}_2)$ and c represents a nonzero $\mathbb{Z}_2$-homology class, the latter equality cannot hold. Hence, L must be the identity.

Therefore, it suffices to prove that $\bar{L}$ is generated by $\{\overline{Y_{ij}} \mid 1\leq i\leq p-1, 1\leq j\leq p, i\neq j\}$. In fact, by our definitions, it suffices to show that $\bar{L}$ is generated by the elementary transformations, $\{E_{ij} \mid 1\leq i\leq p-1, 1\leq j\leq p-1, i\neq j\}$, and the sign change transformations, $\{F_i \mid 1\leq i\leq p-1\}$. This is what we now proceed to demonstrate.

As suggested by our previous comments, our proof requires that we consider matrices with respect to the $\mathbb{Z}$-basis for R_p, $\{c_1,\ldots,c_{p-1}\}$.

We shall denote automorphisms of R_p and their corresponding matrices with respect to this basis by the same symbol. In particular, by our assumptions, $\bar{L}$ is congruent modulo 2 to the identity matrix.

(2) $\bar{L} \equiv I_{p-1}$ (modulo 2).

Moreover, since $\bar{L}$ is invertible over $\mathbb{Z}$, its determinant must be a unit of $\mathbb{Z}$.

(3) $\det(\bar{L}) = \pm 1$.

Given a matrix, M, let k(M) denote the sum of the absolute values of the entries of the first column of M.

(4) $$k(M) = \sum_{i=1}^{p-1} |M(i,1)|.$$

Let $k=k(\bar{L})$. By (2), k is odd. If $k \geq 3$, we apply (3) to conclude that $\bar{L}(j,1)$ is nonzero for some $j \geq 2$. Choosing such a j, we apply (2.10) to compute:

(5) $$k(E_{1j}\bar{L}) = |\bar{L}(1,1)+2\bar{L}(j,1)| + \sum_{i=2}^{p-1} |\bar{L}(i,1)|.$$

In an equivalent form, we conclude:

(6) $$k(E_{1j}\bar{L}) + |\bar{L}(1,1)| = k(\bar{L}) + |\bar{L}(1,1)+2\bar{L}(j,1)|.$$

In similar fashion, we have the following identities:

(7) $$k(E_{1j}^{-1}\bar{L}) + |\bar{L}(1,1)| = k(\bar{L}) + |\bar{L}(1,1)-2\bar{L}(j,1)|,$$

(8) $k(E_{j1}\bar{L})+|\bar{L}(j,1)|=k(\bar{L})+|\bar{L}(j,1)+2\bar{L}(1,1)|$,

(9) $k(E_{j1}^{-1}\bar{L})+|\bar{L}(j,1)|=k(\bar{L})+|\bar{L}(j,1)-2\bar{L}(1,1)|$.

If we apply Lemma 2.1 with $a=\bar{L}(1,1)$ and $b=\bar{L}(j,1)$, we conclude that one of the following inequalities must hold.

(10) (i) $k(E_{1j}\bar{L})<k(\bar{L})$

(ii) $k(E_{1j}^{-1}\bar{L})<k(\bar{L})$

(iii) $k(E_{j1}\bar{L})<k(\bar{L})$

(iv) $k(E_{j1}^{-1}\bar{L})<k(\bar{L})$.

Therefore, by repeated multiplications on the left by elementary transformations, we may reduce to the case where k=1. If k=1, however, $\bar{L}$ must have the following form.

(11) k=1

$$\bar{L} = \begin{pmatrix} \varepsilon & 2m_2 & \cdots & 2m_{p-1} \\ 0 & & & \\ \cdot & & & \\ \cdot & & M & \\ \cdot & & & \\ 0 & & & \end{pmatrix}$$

Here $\varepsilon=\pm 1$, $m_2, \ldots, m_{p-1}$ are integers and M is a $(p-2)\times(p-2)$ integer matrix. In this case, therefore, one easily computes the following product.

(12) k=1

$$F_i\bar{L}E_{12}^{m_2}\cdot\ \cdot\ \cdot E_{1,p-1}^{m_{p-1}} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \cdot & & & \\ \cdot & & \bar{L}_1 & \\ \cdot & & & \\ 0 & & & \end{pmatrix}$$

Here $\bar{L}_1$ is a (p-2)x(p-2) matrix. The theorem follows by an obvious induction.

Remark By our remarks concerning the automorphisms in (2.2), we see that each of these automorphisms is induced by a conjugate of σ by a diffeomorphism which is generated by τ and Ψ. Hence, each of these automorphisms is induced by a crosscap slide. The original statement of Theorem 3 which was formulated in the introduction follows immediately.

As a consequence of our observations at the beginning of the proof, we have identified the group of automorphisms of $H_1(F_p,\mathbb{Z})$ which act trivially on $H_1(F_p,\mathbb{Z}_2)$ with a subgroup of the group of automorphisms of R_p which act trivially on $R_p \otimes \mathbb{Z}_2$. Since there is a natural way of lifting automorphisms of R_p to automorphisms of $H_1(F_p,\mathbb{Z})$, it is easy to see that this subgroup is actually the full group of such automorphisms. On the other hand, this full group is isomorphic, under the obvious identifications, with the congruence subgroup modulo 2 of $GL(p-1,\mathbb{Z})$, $\Gamma_2(p-1,\mathbb{Z})$. Hence, we have the following immediate corollary.

Corollary 2.3 $\Gamma_2(p-1,\mathbb{Z})$ is generated by the squares of standard elementary matrices,

$$E_{ij} \qquad 1 \leq i \leq p-1 \quad 1 \leq j \leq p-1 \quad i \neq j$$

and the sign change transformations,

$$F_1, \ . \ . \ . \ . \ , \ F_{p-1}.$$

Section 3

Finally, we state and prove a more precise formulation of Theorem 1. The notation refers to the discussion in sections 1 and 2.

Theorem 3.1 If L is an automorphism of $H_1(F_p,\mathbb{Z})$ which preserves the associated $\mathbb{Z}_2$-valued intersection pairing, then L is induced by a diffeomorphism of F_p which is generated by the crosscap transposition, τ, the crosscap p-cycle, Ψ, the Dehn twist, δ, and the crosscap slide, σ.

Proof: Let L be an automorphism of $H_1(F_p,\mathbb{Z})$ as above. Let L_2 be the induced automorphism of $H_1(F_p,\mathbb{Z}_2)$. By Theorem 1.1, L_2 is induced by a diffeomorphism, ν, which is generated by τ,Ψ and δ. Let $L'=(\nu^{-1})_*\circ L$ where $(\nu^{-1})_*$ is the action of ν^{-1} on $H_1(F_p,\mathbb{Z})$. Clearly, L' acts trivially on $H_1(F_p,\mathbb{Z}_2)$.

By Theorem 2.2, L' is generated by $\{Y_{ij} | 1\le i\le p-1,\ 1\le j\le p,\ i\ne j\}$. We have previously observed that each transformation, Y_{ij}, is induced by an appropriate conjugate of σ by a diffeomorphism in the group generated by τ and Ψ. Hence, L' is induced by a diffeomorphism of the desired type.

Since $L=\nu_*\circ L'$, L is also induced by a diffeomorphism of the desired type.

Remark This formulation of the result corresponds to the result for orientable surfaces which was obtained from Hua and Reiner's system of generators for $Sp(2g,\mathbb{Z})$ [HR], [Bi]. Here, the symplectic group is replaced by a $\mathbb{Z}_2$-orthogonal group".

Acknowledgement The authors gratefully acknowledge the support of the Max-Planck-Institut für Mathematik, Bonn during the preparation of this paper.

References

[Bi] Birman, J.S., On Siegel's modular group, Math. Ann. 191 (1971), 59-68.

[Bu] Burkhardt, H., Grundzüge einer allgemeinen Systematik der hyperelliptischen Funktionen erster Ordnung, Math. Ann. 35 (1890), 198-296.

[Ch] Chillingworth, D.R.J., A finite set of generators for the homeotopy group of a nonorientable surface, Proc. Cambridge Philos. Soc. 65 (1969), 409-430.

[HR] Hua, L.K. and Reiner, I., On the generators of the symplectic modular group, Trans. Am. Math. Soc. 65 (1949), 415-426.

[L] Lickorish, W.B.R., Homeomorphisms of non-orientable 2-manifolds, Proc. Cambridge Philos. Soc. 59 (1963), 307-317.

[MP] Meeks, W.H. III and Patrusky, J., Representing homology classes by embedded circles on a compact surface, Ill. Jour. Math. 22, no. 2, (1978), 262-269.