JOURNAL OF ALGEBRA 32, 317–327 (1974)

Characters of F_2 Represented in Sp(4, R)

NANCY MAKIESKY ZUMOFF

Department of Mathematics, Oberlin College, Oberlin, Ohio 44074

Communicated by H. J. Ryser

Received June 29, 1973

1. INTRODUCTION.

Given any set of matrix representations of a group, we can derive information about the group and the representations by studying the characters of the representation. In particular, we can consider properties of the character which will hold for all representations in this set. For example, we say that two elements u and v lie in the same character class if they have the same character for all representations in the set.

R. Horowitz [2] has considered all representations of a free group on n free generators in the 2×2 special linear group $\mathrm{SL}(2, K)$ (K an integral domain). He shows that there exists a set of $2^n - 1$ words u_i in F_n such that the character of any word u in F_n can be represented as a polynomial in the characters of the words $u_1, \ldots, u_{2^n-1}$. That is, for every u in F_n there exists a polynomial P in $2^n - 1$ variables such that $\mathrm{tr}(\rho(u)) = P[\mathrm{tr}(\rho(u_i))]$ where ρ is a representation of F_n in $\mathrm{SL}(2, K)$ and where tr denotes the trace of the matrix.

Changing the set of representations considered changes this property. A. V. Marincuk and K. S. Sibirskii [4] show that if one considers representations of F_2 in $\mathrm{GL}(3, \mathbf{R})$, (the general linear group of 3×3 matrices with coefficients in $\mathbf{R}$) then the character of any element can be represented as a polynomial in eleven elements.

We will consider representations of F_2 in the 4×4 symplectic group $\mathrm{Sp}(4, \mathbf{R})$ where $\mathrm{Sp}(4, \mathbf{R})$ is given by the algebraic condition: $M \in \mathrm{Sp}(4, \mathbf{R})$ if M is a real. 4×4 matrix such that $M^T J M$ equals J where M^T is the transpose of M and $J = \left(\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right)$ with 0 and I the 2×2 zero and identity matrices, respectively. Equivalently, $\mathrm{Sp}(4, \mathbf{R})$ can be described as the group of analytic mappings of the generalized upper half plane $\mathscr{H}$ where $\mathscr{H}$ consists of all 2×2 symmetric matrices $\mathscr{Z}$ such that $(1/2i)(\mathscr{Z} - \bar{\mathscr{Z}})$ is positive definite [5].

If we consider 4×4 symplectic representations, we increase the number of words needed whose characters generate all characters: If ρ is a representation of F_2 in SL(2, R) then ρ induces a representation ρ^* in Sp(4, R) by

$$\rho^*(W) = \begin{pmatrix} \rho(W) & 0 \\ 0 & (\rho(W)^T)^{-1} \end{pmatrix}.$$

Then $\operatorname{tr}(\rho^*(W)) = 2\operatorname{tr}(\rho(W))$. Therefore any polynomial relationship between characters in Sp(4, R) must induce relationships in SL(2, **R**). However, since not all symplectic representations are derivable in this way, not all relationships holding in SL(2, **R**) will imply relationships in Sp(4, **R**).

We prove the following analogous theorem for Sp(4, R): There exists a set of twenty words $\{W_1, \ldots, W_{20}\}$ in F_2 such that the character of any word in F_2 is representable as a polynomial in characters of these words.

If two words are conjugate as group elements of F_2 then they lie in the same character class, no matter what set of representations we consider, since conjugate matrices have equal trace. If we consider representations in SL(2, **R**) or Sp(4, **R**), then an element and its inverse lie in the same character class, since if a matrix A lies in either group, $\operatorname{tr} A = \operatorname{tr} A^{-1}$. Therefore if u is in F_2 then the character class of u is the union of at least two conjugacy classes, that of u and of u^{-1}. Horowitz [2] has given certain necessary conditions for words in F_2 to be in the same character class, and these conditions must carry over to representations in Sp(4, **R**). However these symplectic representations distinguish conjugacy classes more sharply than SL(2, **R**) representations. We will give examples of words which lie in the same character class with respect to representations in SL(2, **R**) but lie in distinct character classes with respect to representations in Sp(4, **R**).

Finally, we show that the set $S = \{W_1, \ldots, W_{20}\}$ has a subset $S_1 = \{W_1, \ldots, W_{12}\}$ invariant under automorphisms of F_2 in the following sense: If ϕ is an automorphism of F_2 and if $W \in S_1$ then the trace of $\phi(W)$ is representable by a polynomial in the traces of words in S_1.

In deriving relationships among the traces of symplectic matrices M we construct a "pseudo-symplectic" normal form M_C. This is a matrix similar to M but not necessarily symplectic: $M_C = C^{-1}MC$ where $C^TJC = J$, so C is symplectic if and only if it is real. This form was derived independently from, but is very closely related to, a symplectic normal form constructed by A. Christian in 1967 [1].

2. Normal Forms for Symplectic Matrices

If M is in Sp(4, **R**) then the following properties of M can be shown by elementary methods of linear algebra.

LEMMA 1. *If λ is an eigenvalue for M then $1/\lambda$ and $\bar{\lambda}$ are also eigenvalues for M. All have the same algebraic and geometric multiplicity.*

LEMMA 2. *If M has an eigenvalue λ such that $\lambda \neq \bar{\lambda}$ and $|\lambda| \neq 1$ then M is diagonalizable.*

LEMMA 3. *If $\lambda \neq \pm 1$ and $\epsilon = \pm 1$ are eigenvalues for M then ϵ has algebraic multiplicity* 2.

LEMMA 4. *The eigenvalues for M, listed according to algebraic multiplicity, can be ordered in one of the following ways:*

(a) $\{\lambda, \mu, 1/\lambda, 1/\mu\}$ *where* $\lambda \neq 1/\lambda, \mu \neq 1/\mu, \lambda \neq \mu^{\pm 1}$,
(b) $\{\lambda, \epsilon, 1/\lambda, \epsilon\}$ *where* $\lambda \neq 1/\lambda, \epsilon = \pm 1$,
(c) $\{\lambda, 1/\lambda, 1/\lambda, \lambda\}$ *where* $\lambda \neq 1/\lambda$,
(d) $\{\epsilon, -\epsilon, \epsilon, -\epsilon\}$ *where* $\epsilon = \pm 1$,
(e) $\{\epsilon, \epsilon, \epsilon, \epsilon\}$ *where* $\epsilon = \pm 1$.

THEOREM 2.1. *There exists a matrix C such that $C^T J C = J$ and $C^{-1}MC = M_C$ is one of the following:*

$$\text{(i)} \quad M_C = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix},$$

$$\text{(ii)} \quad M_C = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & \epsilon & 0 & 1 \\ 0 & 0 & 1/\lambda & 0 \\ 0 & 0 & 0 & \epsilon \end{pmatrix},$$

$$\text{(iii)} \quad M_C = \begin{pmatrix} \lambda & 0 & 0 & \lambda \\ 0 & 1/\lambda & 1/\lambda & 0 \\ 0 & 0 & 1/\lambda & 0 \\ 0 & 0 & 0 & \lambda \end{pmatrix},$$

$$\text{(iv)} \quad M_C = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

$$\text{(v)} \quad M_C = \begin{pmatrix} \epsilon I & S \\ 0 & I \end{pmatrix}, \qquad \textit{where} \quad S = S^T,$$

$$\text{(vi)} \quad M_C = \begin{pmatrix} \epsilon & \epsilon & 0 & 0 \\ 0 & \epsilon & \epsilon & -\epsilon \\ 0 & 0 & \epsilon & 0 \\ 0 & 0 & -\epsilon & \epsilon \end{pmatrix}.$$

Proof. If X has column vectors x_i then $X^TJX = ((x_i, Jx_j))$ where $(\ ,\ )$ is the symmetric inner product. If M is symplectic, then $(x_i, Jx_j) = (Mx_i, JMx_j)$.

If the eigenvalues are of type (a), notice that $\lambda_i\lambda_j = 1$ if and only if $(i, j) = (1, 3)$, $(2, 4)$, $(3, 1)$, or $(4, 2)$. Choosing eigenvectors x_i, we find $(x_i, Jx_j) = \lambda_i\lambda_j(x_i, Jx_j)$. Then $(x_i, Jx_j) = 0$ except for those values of (i, j). Since $\det X \neq 0$ and $\det X^TJX = (\det X)^2$, then (x_i, Jx_j) cannot be zero for those remaining (i, j). The x_i can be normalized such that $(x_i, Jx_j) = 1$ for these values. If C has the normalized eigenvectors as its columns, then $C^{-1}MC$ is diagonal and $C^TJC = J$.

For eigenvalues of remaining types we choose x_i as indicated by the possible Jordan canonical forms, rearranged if necessary to give the ordering of eigenvalues indicated by Lemma 4. The argument above can be adapted to shoe $C^{-1}MC$ is of the desired form and $C^TJC = J$. All necessary computations are routine.

3. Basic Relationship for Characters

U, V, W,... will denote 4×4 symplectic matrices. $\operatorname{tr} U$ denotes the trace of U.

Theorem 3.1. $\operatorname{tr} UVW + \operatorname{tr} U^{-1}VW + \operatorname{tr} UV^{-1}W + \operatorname{tr} U^{-1}V^{-1}W + \operatorname{tr} VUW + \operatorname{tr} VU^{-1}W + \operatorname{tr} V^{-1}UW + \operatorname{tr} V^{-1}U^{-1}W = \operatorname{tr} U[\operatorname{tr} VW + \operatorname{tr} V^{-1}W] + \operatorname{tr} V[\operatorname{tr} UW + \operatorname{tr} U^{-1}W] + \operatorname{tr} W[\operatorname{tr} UV + \operatorname{tr} U^{-1}V] - \operatorname{tr} U \operatorname{tr} V \operatorname{tr} W$.

Proof. Assume U is in normal form. V and W are no longer necessarily symplectic, since they may no longer be real, but still satisfy $V^TJV = J$, $W^TJW = J$. In particular, if $V = \left(\begin{smallmatrix} V_1 & V_2 \\ V_3 & V_4 \end{smallmatrix}\right)$, where V_i are 2×2 matrices, then

$$V + V^{-1} = \begin{pmatrix} V_1 + V_4^T & V_2 - V_2^T \\ V_3 - V_3^T & V_4 + V_1^T \end{pmatrix}.$$

If U has form (i), (ii), (iv), or (v) in Theorem 2.1 and U_0 is a diagonal matrix having the same diagonal entries as U then $\operatorname{tr} U = \operatorname{tr} U_0$ and $U + U^{-1} = U_0 + U_0^{-1}$. Therefore in these cases it is sufficient to prove the theorem for U diagonal.

Let $V = (v_{ij})$, $V^{-1} = (\tilde{v}_{ij})$, $W = (w_{ij})$. Note that $v_{ij} + \tilde{v}_{ij} = 0$ if $(i, j) = (1, 3), (2, 4), (3, 1), (4, 2)$. Let U have diagonal entries $\lambda_1, \lambda_2, \lambda_3, \lambda_4$.

Then the left-hand side of the equation is

$$\sum_{1\leqslant i,j\leqslant 4} (\lambda_i + 1/\lambda_i + \lambda_j + 1/\lambda_j)(v_{ij} + \tilde{v}_{ij})\, w_{ji}$$

$$= 2\sum_{i=1}^{4} (\lambda_i + 1/\lambda_i)(v_{ii} + \tilde{v}_{ii})\, w_{ii}$$

$$+ \sum_{i\neq j} (\lambda_i + 1/\lambda_i + \lambda_j + 1/\lambda_j)(v_{ij} + \tilde{v}_{ij})\, w_{ji}\,.$$

If $i \neq j$ then either $\lambda_i + 1/\lambda_i + \lambda_j + 1/\lambda_j = \operatorname{tr} U$ or $v_{ij} + \tilde{v}_{ij} = 0$. Therefore this is equal to

$$2\sum_{i=1}^{4} (\lambda_i + 1/\lambda_i)(v_{ii} + \tilde{v}_{ii})\, w_{ii}$$

$$+ \operatorname{tr} U \left[\sum_{1\leqslant i,j\leqslant 4} (v_{ij} + \tilde{v}_{ij})\, w_{ji} - \sum_{i=1}^{4} (v_{ii} + \tilde{v}_{ii})\, w_{ii} \right].$$

Direct inspection shows this is equal to the right-hand side.

If U has the form (iii) or (vi) then $U + U^{-1} = aI + bX$ where $X = \left(\begin{smallmatrix} 0 & L \\ 0 & 0 \end{smallmatrix}\right)$, $L = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $a = \lambda + 1/\lambda$, $b = \lambda - 1/\lambda$ [form (iii)] or $a = 2$, $b = -1$ [form (vi)]. Since tr is additive, and since the theorem has been proved for diagonal matrices, it is sufficient to prove it for bX. A routine calculation does this.

COROLLARY 3.2. *The following relations hold:*

(a) $$\operatorname{tr} U^2V = -\operatorname{tr} U^{-2}V + \operatorname{tr} U[\operatorname{tr} UV + \operatorname{tr} U^{-1}V] - \tfrac{1}{2}[(\operatorname{tr} U)^2 - \operatorname{tr}(U^2)] \operatorname{tr} V.$$

(b) $$\operatorname{tr} U^n = -\operatorname{tr} U^{n-4} + \operatorname{tr} U[\operatorname{tr} U^{n-1} + \operatorname{tr} U^{n-3}] - \tfrac{1}{2}[(\operatorname{tr} U)^2 - \operatorname{tr}(U^2)] \operatorname{tr} V.$$

(c) $$\begin{aligned} &\operatorname{tr} UVUV + \operatorname{tr} U^{-1}VU^{-1}V \\ &\qquad + 2[\operatorname{tr} UVUV^{-1} + \operatorname{tr} UVU^{-1}V + \operatorname{tr} UVU^{-1}V^{-1}] \\ &= [\operatorname{tr} UV + \operatorname{tr} U^{-1}V]^2 \\ &\quad - \tfrac{1}{2}[(\operatorname{tr} U)^2 - \operatorname{tr}(U^2) - 4][(\operatorname{tr} V)^2 - \operatorname{tr}(V^2) - 4]. \end{aligned}$$

Proof. (a) Apply Theorem 3.1 to U, U, and V.

(b) Apply part (a) with $V = U^{n-2}$.

(c) Apply Theorem 3.1 to U, V, UV then to U, V, $U^{-1}V$.

Add the results and simplify.

4. Polynomial Representation of Characters

Theorem 4.1. *Let $F_2 = \langle a, b \rangle$ be a free group on two free generators. If $u \in F_2$, then the character of u can be represented as a polynomial* $\operatorname{tr} u = \underline{P}(x_1, \ldots, x_{20})$ *where* $x_i = \operatorname{tr} W_i$ *and* $S = \{W_1, \ldots, W_{20}\} = \{a, b, a^2, b^2, ab, ab^{-1}, a^2b, ab^2, a^2b^2, (ab)^2, abab^{-1}, aba^{-1}b, a^2bab^{-1}, ab^2a^{-1}b, a^2b^2ab, a^2b^2ab^{-1}, a^2baba^{-1}b, ab^2abab^{-1}, a^2b^2abab^{-1}, a^2b^2aba^{-1}b\}$. *That is,* $\operatorname{tr}(\rho(u)) = \underline{P}(\operatorname{tr}(\rho(W_i)))$ *for all representations* ρ *of* F_2 *by* 4×4 *symplectic matrices.*

Proof. Let $u = a^{\alpha_1}b^{\beta_1} \cdots a^{\alpha_n}b^{\beta_n}$. Define $L(u) = n$ (call it the L-length of u) and $k(u) =$ number of i such that $|\alpha_i| = 2$ or $|\beta_i| = 2$ (call it the k-length). We write $\underline{P}(\operatorname{tr} U_i) \equiv 0$, where $\underline{P}$ is a polynomial if there exist a polynomial Q and $v_j \in F_2$ such that $\underline{P}(\operatorname{tr} u_i) = Q(\operatorname{tr} v_j)$ and either:

(i) $L(v_j) \leqslant L(u_i)$ for all i, j and $k(v_j) < k(u_i)$ whenever $L(v_j) = L(u_i)$ and $L(u_i)$ minimal,

(ii) $v_j \in S$.

That is, $P(\operatorname{tr} u_i) \equiv 0$ if it can be expressed as a polynomial in traces of words with shorter L-length or equal L-length and shorter k-length. Note that if $\operatorname{tr} u + \operatorname{tr} v \equiv 0$, $\operatorname{tr} u \equiv 0$, and $L(u) \leqslant L(v)$ then $\operatorname{tr} v \equiv 0$.

The proof proceeds through a series of reductions.

Step 1. It is sufficient to consider u freely and cyclically reduced. If $u \neq 1$, then none of its exponents are zero.

Step 2. It is sufficient to consider u containing only 1, -1, or 2 as exponents. If $|\alpha_i| > 2$, $|\beta_j| > 2$, $\alpha_i = -2$ or $\beta_j = -2$ then repeated application of Corollary 3.2(a) yields $\operatorname{tr} u = P(\operatorname{tr} u_i)$ where u_i involves only $\pm 1, 2$ as exponents.

Step 3. $\operatorname{tr}(u_i u_2 u_3 \cdots u_n) + \operatorname{tr}(u_2 u_1 u_3 \cdots u_n) \equiv 0$, where $u_i = a^{\alpha_i}b^{\beta_i}$. This follows from Theorem 3.1 with $U = u_1$, $V = u_2$, $W = u_3 \cdots u_n$.

Step 4. $\operatorname{tr} u \equiv 0$ if $k(u) \geqslant 2$. By the previous steps we may assume $u = u_1 \cdots u_n$ with $k(u_1) = k(u_2) = 2$. Then $u_i = a^2b^\epsilon$, $a^\epsilon b^2$, or a^2b^2 for $i = 1, 2$, where $\epsilon = \pm 1$. Interchanging u_1 and u_2 if necessary, we may further assume u has one of the following forms:

(a) $u = a^2b^\epsilon a^2b^\alpha v$,

(b) $u = a^\alpha b^2 a^\epsilon b^2 v$,

(c) $u = a^2b^2a^2b^2v$,

(d) $u = a^2b^\epsilon a^\delta b^2 v$,

where $\epsilon, \delta = \pm 1$, $\alpha = \pm 1, 2$ and $v = 1$ or $v = u_3 \cdots u_n$. For u of type (a) of (b) consider $w = x^2y^\epsilon x^2y^\alpha z$. Then $w = u$ for $x = a$, $y = b$, u of type (a) and w is conjugate to u for $x = b, y = a$, u of type (b). Applying Theorem 3.1 to x, xy^ϵ, $x^2y^\alpha z$ we find $\operatorname{tr} w \equiv 0$. For of type (c), applying Theorem 3.1 to a, ab^2, a^2b^2v we find $\operatorname{tr} u + \operatorname{tr}(ab^2a^3b^2v) \equiv 0$. Applying Corollary 3.2(a) to $a^2 \cdot ab^2vab^2$ we find $\operatorname{tr}(ab^2a^3b^2v) \equiv 0$ so $\operatorname{tr} u \equiv 0$. Finally, if u is of type (d) then by Corollary 3.2(a) we can consider instead $u = a^{2\delta}b^\epsilon a^\delta b^2 v$. Applying Theorem 3.1 to a^δ, $a^\delta b^\epsilon$, $a^\delta b^2 v$ we obtain $\operatorname{tr} u \equiv 0$.

Step 5. If $L(u) \geqslant 3$ and $u_i = u_j$ for some $i \neq j$ then $\operatorname{tr} u \equiv 0$. By Step 2 we may assume $u_1 = u_2$. Then $u = u_1{}^2 v$ where $v \neq 1$. Applying Corollary 3.2(a) we find $\operatorname{tr} u \equiv 0$.

Step 6. If $L(u) \geqslant 4$ or $L(u) = 3$ and $k(u) = 0$ then $\operatorname{tr} u \equiv 0$. By the preceeding steps we may assume u has at least three L-syllables with 0 k-length, and that $u = u_1u_2u_3v$ where $k(u_1) = k(u_2) = k(u_3) = 0$ and the u_i are distinct. Since $u_i = a^{\epsilon_i}b^{\delta_i}$, ϵ_i, $\delta_i = \pm 1$ we may assume $\epsilon_2 = \epsilon_3$, $\delta_2 = -\delta_3$. Then $\epsilon_1 = -\epsilon_2$ and $\delta_1 = \delta_2$ or δ_3. By interchanging u_2 and u_3 if necessary, assume $\delta_1 = \delta_2$. Then $u = a^{-\epsilon}b^\delta a^\epsilon b^\delta a^\epsilon b^{-\delta}v = a^{-\epsilon}(b^\delta a^\epsilon)^2 b^{-\delta} v$ and by Step 6, $\operatorname{tr} u \equiv 0$.

Step 7. There remain only a finite number of words u to consider. These have the following properties:

(i) u involves only exponents $\pm 1, 2$,
(ii) $L(u) \leqslant 3$ and $k(u) = 1$ if $L(u) = 3$,
(iii) u has no repeated syllables if $L(u) = 3$.

We first make the following observation. If three out of the four words abv, $ab^{-1}v$, $a^{-1}bv$, $a^{-1}b^{-1}v$ are $\equiv 0$ then the fourth is. This is an immediate consequence of Theorem 3.1. We can then show $\operatorname{tr} u \equiv 0$ for all remaining u by applying Theorem 3.1, Corollary 3.2 and this observation and successively reducing words, beginning with $L(u) = 3$, $k(u) = 1$. Straight forward reduction gives the result.

5. INVARIANT SUBSETS OF S

We let $S = \{W_1, \ldots, W_{20}\}$, ordered as in the previous section, and let P denote a polynomial in the indeterminants $x_1, \ldots, x_{20}$. Following Horowitz's terminology, a polynomial P represents a word $u \in F_2$ if $\operatorname{tr}(\rho(u)) = P(\operatorname{tr}(\rho(u_i)))$ for all representations ρ of F_2 by 4×4 symplectic matrices. With this terminology, Theorem 4.1 now states that every word in a and b is represented by such a polynomial P. We say that a polynomial represents 0 if $P(\operatorname{tr}(\rho(u_i)) = 0$ for all ρ.

We can now consider the subset $S_1 \subset S$ consisting of $\{W_1, \ldots, W_{12}\} = \{a, b, a^2, b^2, ab, ab^{-1}, a^2b, ab^2, a^2b^2, (ab)^2, abab^{-1}, aba^{-1}b\}$, and let R_1 be the ring of polynomials with rational coefficients in the indeterminants $X_1, \ldots, X_{12}$. Then $P \in R_1$ represents $u \in F_2$ if

$$\operatorname{tr}(\rho(u)) = P(\operatorname{tr}(\rho(W_1)), \ldots, \operatorname{tr}(\rho(W_{12}))).$$

THEOREM. *If σ is an automorphism of F_2, and if $W_i \in S_1$ then $\sigma(W_i)$ can be represented by a polynomial in R_1.*

Proof. Let Φ_2 be the group of automorphisms of F_2. Then Φ_2 is generated by the automorphisms ϕ_1, ϕ_2, ϕ_3 where

$$\begin{aligned} \phi_1(a) &= a^{-1}, & \phi_1(b) &= b, \\ \phi_2(a) &= b, & \phi_2(b) &= a, \\ \phi_3(a) &= ab, & \phi_3(b) &= b, \end{aligned}$$

(see [3]).

Therefore it is sufficient to prove the theorem for the generating automorphisms ϕ_1, ϕ_2, ϕ_3. We write $W \sim W^1$ if W is conjugate to W^1. Then we have the following:

i	W_i	$\phi_1(W_i)$	$\phi_2(W_i)$	$\phi_3(W_i)$
1	a	$a^{-1} = W_1^{-1}$	$b = W_2$	$ab = W_5$
2	b	$b = W_2$	$a = W_1$	$b = W_2$
3	a^2	$a^{-2} = W_3^{-1}$	$b^2 = W_4$	$(ab)^2 = W_{10}$
4	b^2	$b^2 = W_4$	$a^2 = W_3$	$b^2 = W_4$
5	ab	$a^{-1}b \sim W_6^{-1}$	$ba \sim W_5$	$ab^2 = W_8$
6	ab^{-1}	$a^{-1}b^{-1} \sim W_5^{-1}$	$ba^{-1} \sim W_6^{-1}$	$a = W_1$
7	a^2b	$a^{-2}b$	$b^2a \sim W_8$	$abab^2$
8	ab^2	$a^{-1}b^2$	$ba^2 \sim W_7$	ab^3
9	a^2b^2	$a^{-2}b^2$	$b^2a^2 \sim W_9$	$abab^3$
10	$(ab)^2$	$(a^{-1}b)^2$	$(ba)^2 \sim W_{10}$	$(ab^2)^2$
11	$abab^{-1}$	$a^{-1}ba^{-1}b^{-1} \sim W_{11}^{-1}$	$baba^{-1} \sim W_{12}$	$ab^2a \sim W_9$
12	$aba^{-1}b$	$a^{-1}bab \sim W_{12}$	$bab^{-1}a \sim W_{11}$	$aba^{-1}b = W_{12}$

If $\phi = \phi_1, \phi_2$, or ϕ_3 and $\phi(W_i) \sim W^{j\pm1}$ then $\operatorname{tr}(\phi(W_i)) = \operatorname{tr} W_j$ and $\phi(W_i)$ is representable by the polynomial X_j in R. Therefore, we need only consider those ϕ_i such that $\phi(W_i) \not\sim W_j^{\pm1}$. $\phi_1(W_i)$ and $\phi_3(W_j)$ are representable by polynomials in $X_1, \ldots, X_{12}$ for $i = 7, 8, 9, 10$; $j = 7, 8, 9$ as immediate consequences of Corollary 3.2. For $\phi_3(W_{10})$ it is necessary to apply Corollary 3.2 to get $\mathrm{tr}(\phi_3(W_{10}) = \underline{P}(\mathrm{tr}\,\phi_i(W_j))$ where $\phi_i(W_j) \neq \phi_3(W_{10})$, then use the preceeding results.

Remark. By inspection it can be shown that S_1 is the only nonempty subset of S with this property. That is, $W \in S$, let $S(W)$ be the smallest subset of S containing W such that if $W_i \in S(W)$ then

$$\mathrm{tr}(\phi(W_i)) = P(\mathrm{tr}(\phi(W_j))), \qquad \text{where} \quad W_j \in S(W).$$

Then

$$S(W) = \begin{cases} S_1 & \text{if} \quad W \in S_1, \\ S & \text{if} \quad W \notin S_1. \end{cases}$$

6. Relationship of Character Classes to Conjugacy Classes

We say that two elements u and v in F_2 have the same character if $\mathrm{tr}(\rho(u)) = \mathrm{tr}(\rho(v))$ for all 4 × 4 symplectic representations ρ. Clearly if v is conjugate to u or u^{-1} (as an element of F_2) then $\mathrm{tr}\,u = \mathrm{tr}\,v$. We will call the set of all elements of F_2 having the same character as u the character class of u. Then the character class of u is the union of at least two conjugacy classes, namely that of u and of u^{-1}.

Lemma 6.1. *If u and v have the same character, then* $\mathrm{tr}(\rho(u)) = \mathrm{tr}(\rho(v))$ *for all representations ρ of F_2 by matrices in the* 2×2 *special linear group* SL(2, $\mathbf{R}$).

Proof. Let ρ be any such representation. We define a symplectic representation ρ^* by

$$\rho^*(W) = \begin{pmatrix} \rho(W) & 0 \\ 0 & (\rho(W)^T)^{-1} \end{pmatrix},$$

$\rho^*(W)$ is a symplectic matrix and ρ^* is a homomorphism. (We use here the fact that ρ is a homomorphism and that $(A^T)^{-1}(B^T)^{-1} = ((AB)^T)^{-1}$.) Then $\mathrm{tr}(\rho^*(W)) = \mathrm{tr}\,\rho(W) + \mathrm{tr}(\rho(W)^T)^{-1} = 2\,\mathrm{tr}\,\rho(W)$. Therefore if u and v have the same character, $\mathrm{tr}\,\rho(u) = \mathrm{tr}\,\rho(v)$ for all representations of F_2 by matrices. in SL(2, $\mathbf{R}$).

R. Horowitz [2] derives certain necessary conditions for u and v to have the same character in SL(2, $\mathbf{R}$). By Lemma 6.1 these conditions are necessary for u and v to have the same character in the symplectic group. In particular, Horowitz shows

(1) the character class of a^m is the union of the conjugacy classes of a^m and a^{-m},

(2) the character class of any power m of a primitive element c in F_2 is the union of the conjugacy classes of c^m and c^{-m},

(3) the character class of $a^l b^m$ is the union of the conjugacy classes of $a^m b^l$ and $b^{-l} a^{-m}$.

These results must also be true in the 4×4 symplectic group according to Lemma 6.1.

In SL(2, **R**), Horowitz shows this is the best possible result. That is, it is no longer true that the character class of a word consists only of the conjugacy classes of the word and its inverse when one takes cyclically reduced words of at least four syllables.

In particular, he shows that given any even number r ($r \neq 0$) and any positive odd number s, the words

$$W(r, s) = a^{-1} b^r a \cdot b^s a^{-1} b^r a$$

and

$$W(r, s) = a b^r a^{-1} b^s a \cdot b^r a^{-1}$$

have the same character in SL(2, **R**) but are not conjugate. Therefore for any r, s as above, the character class of $b^{2r} a b^s a^{-1}$ is the union of at least four conjugacy classes, namely the conjugacy classes of $b^{2r} a b^s a^{-1}$, $b^{2r} a^{-1} b^s a$, $a b^{-s} a^{-1} b^{-2r}$, and $a^{-1} b^{-s} a b^{2r}$. This is no longer true if we use 4×4 symplectic representations. In this case the character class of $b^{2r} a b^s a^{-1}$ in SL(2, **R**) will split into at least two character classes.

To show this, will construct a representation ρ such that

$$\operatorname{tr}(\rho(W(r, s))) \neq \operatorname{tr}(\rho(V(r, s))).$$

We define

$$\rho(a) = \begin{pmatrix} 0 & 1 & 0 & -1 \\ t & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1/t & 0 \end{pmatrix},$$

$$\rho(b) = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 \\ 0 & 0 & 1/\lambda & 0 \\ 0 & 0 & 0 & 1/\mu \end{pmatrix}.$$

Then

$$\operatorname{tr} \rho(W(r, s)) = \lambda^{2r}/\mu^s + \mu^{2r}\lambda^s + \mu^s/\lambda^{2r} + 1/\mu^{2r}\lambda^s,$$
$$\operatorname{tr} \rho(V(r, s)) = \lambda^{2r}\mu^s + \mu^{2r}/\lambda^s + 1/\lambda^{2r}\mu^s + \lambda^s/\mu^{2r}.$$

Then $\operatorname{tr}(\rho(W(r, s))) = \operatorname{tr}(\rho(V(r, s)))$ if and only if

$$(\lambda^{2r} - 1/\lambda^{2r})(\mu^s - 1/\mu^s) = (\lambda^s - 1/\lambda^s)(\mu^{2r} - 1/\mu^{2r}).$$

In order for this to be true for such r, s and arbitrary real numbers λ, μ (nonzero), the polynomial

$$X^s Y^{2r}(X^{4r} - 1)(Y^{2s} - 1) - X^{2r} Y^s (X^{2s} - 1)(Y^{4r} - 1)$$

must be identically zero. This polynomial has $X^{4r+s}Y^{2r+2s} - X^{2r+2s}Y^{4r+s}$ as its highest degree term and would vanish only if $2r = s$. But s is odd, so this is impossible. Therefore $\operatorname{tr}(\rho(W(r, s))) \neq \operatorname{tr}(\rho(V(r, s)))$.

ACKNOWLEDGMENT

The author thanks Professor Wilhelm Magnus for suggesting this topic and for his interest, encouragement, and advice.

REFERENCES

1. U. CHRISTIAN, A reduction theory for symplectic matrices, *Math. Z.* **101** (1967), 213–244.
2. R. HOROWITZ, Characters of free groups represented in the two dimensional special linear group, *Comm. Pure Appl. Math.* **25** (1972), 635–650.
3. W. MAGNUS, A. KARASS, AND D. SOLITAR, "Combinatorial Group Theory," Interscience, New York, 1966.
4. A. V. MARINCUK AND K. S. SIBIRSKII, Minimal polynomial bases of the affine invariants of three 3 × 3 matrices, *Mat. Issled.* **4** (1969), 46–56.
5. C. L. SIEGEL, "Symplectic Geometry," Academic Press, New York, 1964.