J. Austral. Math. Soc. (*Series A*) **28** (1979), 401–412

TRACE POLYNOMIALS OF WORDS IN SPECIAL LINEAR GROUPS

J. B. SOUTHCOTT

(Received 15 January 1979)

Communicated by H. Lausch

Abstract

If w is a group word in n variables, $x_1, \ldots, x_n$, then R. Horowitz has proved that under an arbitrary mapping of these variables into a two-dimensional special linear group, the trace of the image of w can be expressed as a polynomial with integer coefficients in traces of the images of $2^n - 1$ products of the form $x_{\sigma 1} x_{\sigma 2} \ldots x_{\sigma m}$, $1 \leqslant \sigma_1 < \sigma_2 < \ldots < \sigma_m \leqslant n$. A refinement of this result is proved which shows that such trace polynomials fall into 2^n classes corresponding to a division of n-variable words into 2^n classes. There is also a discussion of conditions which two words must satisfy if their images have the same trace for any mapping of their variables into a two-dimensional special linear group over a ring of characteristic zero.

1980 Mathematics subject classification (*Amer. Math. Soc.*): 20 G 99.

1. Preliminaries

To facilitate the discussion in subsequent sections the terminology to be used will be established immediately.

We shall define a *free group* on a given set of generators as the set of all freely reduced words in the generators, with multiplication of two elements defined by juxtaposition and free reduction of the result.

4 primitive element is an element of a free group which is in a set of free generators for the group.

A word is an element of a free group $X = \text{gp}(x_1, x_2, \ldots)$ of countable rank. In particular, note that a word is always freely reduced. As a rule, the letters occurring in a word may be assumed without loss of generality to be $x_1, \ldots, x_n$, that is, to lie in the group $X_n = gp(x_1, \ldots, x_n)$. In this situation the letters x_i are referred to as variables and the word is referred to as a *word in n variables*.

If A is a group, α a mapping of the free generators of X into A, then the image of a word w under the corresponding homomorphism $\alpha\colon X \to A$ is called a *value* of the word w in A.

https://doi.org/10.1017/S1446788700012544 Published online by Cambridge University Press

The *verbal subgroup* of a group A corresponding to a set of words is the subgroup generated by all values in A of words in the set. In the case where the set consists of a single word w, the verbal subgroup is denoted by $w(A)$.

The length of a word $w = x_{\nu_1}^{\alpha_1} x_{\nu_2}^{\alpha_2} \dots x_{\nu_r}^{\alpha_r}$ is defined as

$$l(w) = \sum_{i=1}^{r} |\alpha_i|.$$

The *exponent sum* on generator x_i in the word w is the length of the image of w under the projection $X \to X$ induced by the mapping of the generators $x_i \to x_i$, $x_j \to 1, j \neq i$.

The symmetric differences of sets μ and ν, that is the set of elements which are in μ or ν but not in both, will be denoted by $\mu \oplus \nu$.

This paper contains results from the author's PhD thesis (University of Queensland, 1975) written under the supervision of Dr Sheila Oates Macdonald. The author wishes to thank Dr Macdonald for her advice and encouragement.

2. Background

Let K be a commutative ring with identity. Then $SL(2, K)$ denotes the two-dimensional special linear group over K, that is the group of all two-by-two matrices of determinant one with entries from K. $PSL(2, K)$ denotes the quotient group of $SL(2, K)$ modulo its centre. If we write $SL(2, q)$ this indicates that the ring is the finite field of order q.

If w is a word in n variables, then the trace of w, denoted by $\operatorname{tr} w$ is defined to be the trace of the value of w under an arbitrary mapping of its variables into $SL(2, K)$. To say that two words u and v have the same trace, $\operatorname{tr} u = \operatorname{tr} v$, means that their traces are equal for all mappings of their variables into $SL(2, K)$. The question of when two words have the same trace in this sense is fundamental and is discussed in Horowitz (1972) and Section 6 below.

THEOREM 2.1. (Horowitz (1972) Theorem 3.1). *If w is a word in n variables, then the trace of w can be expressed as a polynomial with integer coefficients in the $2^n - 1$ traces of the form*

$$t_\sigma = t_{\sigma_1 \cdots \sigma_m} = \operatorname{tr} x_{\sigma_1} \dots x_{\sigma_m}, \quad 1 \leqslant \sigma_1 < \dots < \sigma_m \leqslant n.$$

under an arbitrary mapping of the variables into $SL(2, K)$.

Polynomials which arise in this way will be referred to as *trace polynomials*.

The main result of this paper is Thoerem 4.1. This is a refinement of Theorem 2.1, and gives strong restrictions on the structure of trace polynomials. It is also a generalization of Theorem 5.2.2 of Cossey, Macdonald and Street (1970) which

https://doi.org/10.1017/S1446788700012544 Published online by Cambridge University Press

holds only for two-variable words. The two-variable form has been applied there and in Southcott (1974a, b) to obtain two-variable laws which hold in certain $PSL(2,q)$.

3. Basic trace identities

Given words, u, v, w it is well known that $\operatorname{tr} u = \operatorname{tr} u^v$, and if e is the empty word, we have (Cossey, Macdonald and Street (1970) 5.2.1, Horowitz (1972) 2.2,2.4)

(3.1) (1) $\operatorname{tr} e = 2$,

(2) $\operatorname{tr} uv = \operatorname{tr} u \operatorname{tr} v - \operatorname{tr} uv^{-1}$.

From these relations follow (Horowitz (1972) 2.1,2.3)

(3) $\operatorname{tr} u^{-1} = \operatorname{tr} u$,

(4) $\operatorname{tr} uvw = \operatorname{tr} u \operatorname{tr} vw + \operatorname{tr} v \operatorname{tr} uw + \operatorname{tr} w \operatorname{tr} uv - \operatorname{tr} u \operatorname{tr} v \operatorname{tr} w - \operatorname{tr} uwv$.

The derivation of 3.1 (4) illustrates many of the techniques of manipulating expressions involving traces of words. Using identities 3.1 (1)–(3) and conjugation as necessary, we have

$$\begin{aligned}\operatorname{tr} uvw &= \operatorname{tr} u \operatorname{tr} vw - \operatorname{tr} uw^{-1}v^{-1},\\ \operatorname{tr} uw^{-1}v^{-1} &= \operatorname{tr} uw^{-1} \operatorname{tr} v - \operatorname{tr} uw^{-1}v,\\ \operatorname{tr} uw^{-1} &= \operatorname{tr} u \operatorname{tr} w - \operatorname{tr} uw,\end{aligned}$$

and

$$\begin{aligned}\operatorname{tr} uw^{-1}v &= \operatorname{tr} vuw^{-1}\\ &= \operatorname{tr} vu \operatorname{tr} w - \operatorname{tr} vuw\\ &= \operatorname{tr} uv \operatorname{tr} w - \operatorname{tr} uwv.\end{aligned}$$

Hence

$$\operatorname{tr} uvw = \operatorname{tr} u \operatorname{tr} vw + \operatorname{tr} v \operatorname{tr} uw + \operatorname{tr} w \operatorname{tr} uv - \operatorname{tr} u \operatorname{tr} v \operatorname{tr} w - \operatorname{tr} uwv.$$

4. Definitions and main theorem

The following material establishes the framework for the statement and proof of Theorem 4.1.

Words in n variables can be divided into 2^n classes, namely the cosets of the verbal subgroup $x^2(X_n)$. Each class of words will be denoted in the form W_λ where λ is a subset of $\{1, 2, \ldots, n\}$.

DEFINITION. Suppose $w \in X_n$. Then $w \in W_\varnothing$ if the exponent sum is even on each generator, and $w \in W_\mathrm{v} = W_{\{\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_m\}}$ if the exponent sum in w is odd on each of the generators $x_{\mathrm{v}_1}, x_{\mathrm{v}_2}, \ldots, x_{\mathrm{v}_m}$ and even on each other generator.

https://doi.org/10.1017/S1446788700012544 Published online by Cambridge University Press

Reçall from Theorem 2.1 that the trace of any word in n variables is a polynomial in the 2^n-1 traces t_σ, σ a non-empty subset of $\{1,2,\dots,n\}$. We shall define a class of operators which act on polynomials in these traces. Each operator will be denoted in the form S_λ where λ is a subset of $\{1,2,\dots,n\}$.

Definition. For each trace t_σ, S_λ is defined by $S_\lambda(t_\sigma)=(-1)^{|\lambda\cap\sigma|}t_\sigma$, and for polynomials, S_λ is defined recursively by

$$S_\lambda(ct_{\zeta_1}t_{\zeta_2}\dots t_{\zeta_s})=c\,S_\lambda(t_{\zeta_1})\,S_\lambda(t_{\zeta_2})\dots S_\lambda(t_{\zeta_s}),$$

where c is a constant, and if $g_1,g_2,\dots,g_r$ are monomials in the variables t_σ

$$S_\lambda(g_1+g_2+\dots+g_r)=S_\lambda(g_1)+S_\lambda(g_2)+\dots+S_\lambda(g_r).$$

These operators will now be used to define classes of polynomials in the variables t_σ; each class will be denoted in the form E_ρ where ρ is a subset of $\{1,2,\dots,n\}$.

Definition. A polynomial f in the 2^n-1 variables t_σ lies in the class E_ρ if for all λ

$$S_\lambda(f)=(-1)^{|\lambda\cap\rho|}f.$$

Note that the sum of two polynomials in E_ρ, for some fixed ρ, also lies in E_ρ.

From the definition of the classes W_λ as cosets of $x^2(X_n)$ it is clear that they form a group under coset multiplication isomorphic to C_2^n, elementary abelian of exponent two on n generators.

The classes of polynomials E_ρ also form a group isomorphic to C_2^n with identity $E_\varnothing$ and multiplication defined by $E_\eta E_\zeta=E_\tau$ where $\tau=\eta\oplus\zeta$. The class E_τ contains all polynomials of the form fg, $f\in E_\eta$, $g\in E_\zeta$.

It may also be noted that the operators S_λ form a group under composition, isomorphic to C_2^n, with identity $S_\varnothing$, generated by $S_1,\dots,S_n$.

Theorem 4.1. *Suppose w is a word in n variables. Then under an arbitrary mapping of the variables into $SL(2,K)$,* $\operatorname{tr} w$ *may be expressed as a polynomial with integer coefficients in the* 2^n-1 *traces t_σ and, moreover,* $\operatorname{tr} w\in E_\lambda$ *if and only if $w\in W_\lambda$.*

Proof. The 'only if' follows easily by *reductio ad absurdum* once we have proved that $w\in W_\lambda$ implies $\operatorname{tr} w\in E_\lambda$.

The proof that $w\in W_\lambda$ implies $\operatorname{tr} w\in E_\lambda$ will be by induction on word length of w. The theorem is true for words of the form 1 and

$$x_{\nu_1}x_{\nu_2}\dots x_{\nu_m},\quad 1\leqslant \nu_1<\nu_2<\dots<\nu_m\leqslant n.$$

Note that if it has been established that $u\in W_\eta$, $\operatorname{tr} u\in E_\eta$ and $v\in W_\zeta$, $\operatorname{tr} v\in E_\zeta$, then $uv\in W_\eta W_\zeta=W_{\eta\oplus\zeta}$ and the expression $\operatorname{tr} u\operatorname{tr} v$ lies in $E_\eta E_\zeta=E_{\eta\oplus\zeta}$.

https://doi.org/10.1017/S1446788700012544 Published online by Cambridge University Press

Suppose that the theorem is true for all words of length less than k, and let $w \in W_\lambda$ be a word of length k. The rest of the proof establishes that it is always possible to express $\operatorname{tr} w$ in terms of traces of words of length less than k, or of known trace, and hence, by induction, that $\operatorname{tr} w$ is a polynomial in the traces t_σ and belongs to E_λ.

The proof may be divided into three cases.

Case 1: $w = x_{\nu_1}^{\alpha_1} x_{\nu_2}^{\alpha_2} \dots x_{\nu_s}^{\alpha_s}$ where for some q, $1 \leqslant q \leqslant s$, $|\alpha_q| \geqslant 2$.

We may assume that $q = s$ and that $\alpha_s \geqslant 2$, since w can be transformed into a word of this form having the same trace by conjunction and inversion if necessary. Then

$$\operatorname{tr} w = \operatorname{tr} w x_{\nu_s}^{-1} \operatorname{tr} x_{\nu_s} - \operatorname{tr} w x_{\nu_s}^{-2}.$$

All words appearing on the right-hand side are of length less than k, so by the induction hypothesis the theorem holds for their traces, and $\operatorname{tr} w \in E_\lambda$.

Case 2: $w = x_{\nu_1}^{\alpha_1} x_{\nu_2}^{\alpha_2} \dots x_{\nu_k}^{\alpha_k}$, where $|\alpha_q| = 1$, $1 \leqslant q \leqslant k$, and the x_{ν_q} are not all distinct. We may consider two subcases.

(1) $w = uv$, where $u = x_{\nu_1}^{\alpha_1} \dots x_{\nu_i}$, $v = x_{\nu_{i+1}}^{\alpha_{i+1}} \dots x_{\nu_k}$ and $\nu_i = \nu_k$.
Then $\operatorname{tr} w = \operatorname{tr} u \operatorname{tr} v - \operatorname{tr} uv^{-1}$. All words on the right-hand side are of length less than k, so by the induction hypothesis the theorem holds for their traces, and $\operatorname{tr} w \in E_\lambda$.

(2) $w = uv$ where $u = x_{\nu_1}^{\alpha_1} \dots x_{\nu_i}^{\alpha_i}$, $v = x_{\nu_{i+1}}^{\alpha_{i+1}} \dots x_{\nu_k}^{-1}$, $\nu_i = \nu_k$.
Then $\operatorname{tr} w = \operatorname{tr} u \operatorname{tr} v - \operatorname{tr} uv^{-1}$. The length of uv^{-1} is not greater than k, hence by Case 1, $\operatorname{tr} uv^{-1} \in E_\lambda$, and it follows that $\operatorname{tr} w \in E_\lambda$.

Case 3: $w = x_{\nu_1}^{\alpha_1} \dots x_{\nu_k}^{\alpha_k}$, $|\alpha_q| = 1$, $1 \leqslant q \leqslant k$, all x_{ν_q} distinct. Either all $\alpha_q = 1$ or we may assume $\alpha_k = -1$. Then $\operatorname{tr} w = \operatorname{tr} w x_{\nu_k} \operatorname{tr} x_{\nu_k}^{-1} - \operatorname{tr} w x_{\nu_k}^2$. The lengths of $w x_{\nu_k}$ and $x_{\nu_k}^{-1}$ are less than k, hence by repeated applications of this procedure, $\operatorname{tr} w$ can be expressed as a sum of terms in E_λ and $\pm \operatorname{tr} x_{\nu_1} x_{\nu_2} \dots x_{\nu_k}$. We must prove that $\operatorname{tr} x_{\nu_1} \dots x_{\nu_k} \in E_\lambda$.

If $1 \leqslant \nu_1 < \nu_2 < \dots < \nu_k \leqslant n$, the result is true. Otherwise, suppose

$$1 \leqslant \nu_i < \nu_j < \dots < \nu_l \leqslant n$$

and write $x_{\nu_1} \dots x_{\nu_k} = a x_{\nu_i} b$. If a or b is empty x_{ν_i} can be moved immediately to the end of the word. Otherwise, by 3.1(4)

$$\operatorname{tr} a x_{\nu_i} b = \operatorname{tr} a \operatorname{tr} x_{\nu_i} b + \operatorname{tr} x_{\nu_i} \operatorname{tr} ab + \operatorname{tr} b \operatorname{tr} a x_{\nu_i} - \operatorname{tr} a \operatorname{tr} x_{\nu_i} \operatorname{tr} b - \operatorname{tr} ab x_{\nu_i}.$$

By repeated application of this procedure $x_{\nu_i}, x_{\nu_j}, \dots, x_{\nu_l}$ can be moved to the end of the word in turn, and $\operatorname{tr} x_{\nu_1} \dots x_{\nu_k}$ can be expressed as a sum of terms in E_λ and $\operatorname{tr} x_{\nu_i} x_{\nu_j} \dots x_{\nu_l}$, which is known and is in E_λ. Hence $\operatorname{tr} w \in E_\lambda$.

https://doi.org/10.1017/S1446788700012544 Published online by Cambridge University Press

5. Examples

In this section we shall consider several examples involving three-variable words. For a polynomial f in the seven variables $t_1, t_2, t_3, t_{12}, t_{13}, t_{23}, t_{123}$ the definitions of the operators S_λ are

$$\begin{aligned}
S_\varnothing(f) &= f,\\
S_1(f) &= f(-t_1, t_2, t_3, -t_{12}, -t_{13}, t_{23}, -t_{123}),\\
S_2(f) &= f(t_1, -t_2, t_3, -t_{12}, t_{13}, -t_{23}, -t_{123}),\\
S_3(f) &= f(t_1, t_2, -t_3, t_{12}, -t_{13}, -t_{23}, -t_{123}),\\
S_{12}(f) &= f(-t_1, -t_2, t_3, t_{12}, -t_{13}, -t_{23}, t_{123}),\\
S_{13}(f) &= f(-t_1, t_2, -t_3, -t_{12}, t_{13}, -t_{23}, t_{123}),\\
S_{23}(f) &= f(t_1, -t_2, -t_3, -t_{12}, -t_{13}, t_{23}, t_{123}),\\
S_{123}(f) &= f(-t_1, -t_2, -t_3, t_{12}, t_{13}, t_{23}, -t_{123}).
\end{aligned}$$

Trace polynomials of words in three or more variables are not necessarily unique. For example,

$$\begin{aligned}
\operatorname{tr} x_1^2 x_2 x_3^2 x_2 &= \operatorname{tr} x_1^2 x_2 \operatorname{tr} x_3^2 x_2 - \operatorname{tr} x_1^2 x_3^{-2},\\
\operatorname{tr} x_1^2 x_2 &= t_1 t_{12} - t_2,\\
\operatorname{tr} x_3^2 x_2 &= t_3 t_{23} - t_2,
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{tr} x_1^2 x_3^{-2} &= t_3 \operatorname{tr} x_1^2 x_3^{-1} - \operatorname{tr} x_1^2\\
&= t_3(t_1 \operatorname{tr} x_1 x_3^{-1} - t_3) - t_1^2 + 2\\
&= t_1 t_3(t_1 t_3 - t_{13}) - t_3^2 - t_1^2 + 2.
\end{aligned}$$

Hence

$$\operatorname{tr} x_1^2 x_2 x_3^2 x_2 = t_1^2 + t_2^2 + t_3^2 - t_1 t_2 t_{12} + t_1 t_3 t_{13} - t_2 t_3 t_{23} + t_1 t_3 t_{12} t_{23} - t_1^2 t_3^2 - 2.$$

But

$$\begin{aligned}
\operatorname{tr} x_1^2 x_2 x_3^2 x_2 &= \operatorname{tr} x_1 x_2 x_3 x_3 x_2 x_1\\
&= t_{123} \operatorname{tr} x_3 x_2 x_1 - \operatorname{tr} x_1 x_2 x_3 x_1^{-1} x_2^{-1} x_2^{-1},
\end{aligned}$$

and partitioning into subwords of length two and applying 3.1(4) we have

$$\begin{aligned}
\operatorname{tr} x_1 x_2 x_3 x_1^{-1} x_2^{-1} x_3^{-1} &= \operatorname{tr} x_1 x_2 \operatorname{tr} x_3 x_1^{-1} x_2^{-1} x_3^{-1}\\
&\quad + \operatorname{tr} x_3 x_1^{-1} \operatorname{tr} x_1 x_2 x_2^{-1} x_3^{-1}\\
&\quad + \operatorname{tr} x_2^{-1} x_3^{-1} \operatorname{tr} x_1 x_2 x_3 x_1^{-1}\\
&\quad - \operatorname{tr} x_1 x_2 \operatorname{tr} x_3 x_1^{-1} \operatorname{tr} x_2^{-1} x_3^{-1}\\
&\quad - \operatorname{tr} x_1 x_2 x_2^{-1} x_3^{-1} x_3 x_1^{-1}\\
&= t_{12}^2 + t_1^2 t_2^3 + t_{13}^2 - 2t_1 t_3 t_{13} + t_{23}^2\\
&\quad - t_1 t_3 t_{12} t_{23} + t_{12} t_{13} t_{23} - 2.
\end{aligned}$$

https://doi.org/10.1017/S1446788700012544 Published online by Cambridge University Press

Hence

$$\begin{aligned}\operatorname{tr} x_1^2 x_2 x_3^2 x_2 = {} & t_{123}(t_1 t_{23}+t_2 t_{13}+t_3 t_{12}-t_1 t_2 t_3-t_{123}) \\ & -t_{12}^2-t_{13}^2-t_{23}^2-t_1^2 t_3^2+2t_1 t_3 t_{13}+t_1 t_3 t_{12} t_{23}-t_{12} t_{13} t_{23}+2.\end{aligned}$$

Note that $x_1^2 x_2 x_3^2 x_2 \in W_\varnothing$ and the trace polynomials calculated for it are in $E_\varnothing$.

The known examples of non-uniqueness are not derived using the algorithm given in the proof of Theorem 4.1. Clearly, that algorithm could be modified so that there are no arbitrary choices involved, and in that case there would be a canonical trace polynomial corresponding to each n-variable word. But even given such a modification of the algorithm, the possibility that two words have the same trace but different canonical forms for their trace polynomials cannot be ruled out.

6. Words with the same trace

In this section we shall look at the problem of determining when two words have the same trace in the sense defined in Section 3, that is, both have the same trace under an arbitrary mapping of their variables into some $SL(2,K)$.

One aspect of this problem is to find in the polynomial ring in 2^n-1 variables over the integers, the ideal which is identically zero when the variables are taken as the traces t_σ. This ideal will be denoted by $I_n(K)$.

THEOREM 6.1 (Horowitz (1972) Theorem 4.1). *If K is a commutative ring of characteristic zero, then $I_1(K)=I_2(K)=\{0\}$, the zero ideal; that is, the trace of any word in one or two variables is a unique polynomial in the traces t_1, t_2 and t_{12}.*

THEOREM 6.2 (Horowitz (1972) Theorem 4.3). *Let K be a ring which contains the rational numbers as a subring. Then $I_3(K)$ is the ideal generated by the polynomial*

$$\begin{aligned}t_1^2+t_2^2+t_3^2+t_{12}^2+t_{13}^2+t_{23}^2-t_1 t_2 t_{12}-t_1 t_3 t_{13}-t_2 t_3 t_{23}+t_{12} t_{13} t_{23}-4 \\ -t_{123}(t_1 t_{23}+t_2 t_{13}+t_3 t_{12}-t_1 t_2 t_3-t_{123}).\end{aligned}$$

I conjecture that this result holds when K is any ring of characteristic zero.

Whittemore (1973) shows that for $n \geqslant 4$, $I_n(K)$ is not principal.

The other aspect of the problem of determining when two words have the same trace is finding group theoretic conditions which such words must satisfy. To simplify the discussion, only cases where traces lie in a ring of characteristic zero will be considered.

For words u and v in any number of variables, we have the condition: *If v is conjugate to u or u^{-1} then* $\operatorname{tr} u = \operatorname{tr} v$.

For two-variable words we can say more than this. If

$$u = x_{v_1}^{\alpha_1} x_{v_2}^{\alpha_2} \dots x_{v_m}^{\alpha_m}, \quad v_i \neq v_{i+1}, \quad 1 \leqslant i < m,$$

then the *reversal of* u, denoted $\operatorname{rev} u$, is the word $x_{\nu_m}^{\alpha_m} \dots x_{\nu_2}^{\alpha_2} x_{\nu_1}^{\alpha_1}$, and the family of syllables of u is $\{x_{\nu_1}^{\alpha_1}, x_{\nu_2}^{\alpha_2}, \dots, x_{\nu_m}^{\alpha_m}\}$.

THEOREM 6.3. *Let u and v be two-variable words. If v is conjugate to u or u^{-1} or the reversal of u or u^{-1} then* $\operatorname{tr} u = \operatorname{tr} v$.

PROOF. It is sufficient to show that $\operatorname{tr} u = \operatorname{tr}(\operatorname{rev} u)$ for any two-variable word u. Without loss of generality, assume $u = x^{\alpha_1} y^{\beta_1} x^{\alpha_2} y^{\beta_2} \dots x^{\alpha_m} y^{\beta_m}$. If u contains only one or two syllables, the theorem is true since $\operatorname{rev} u$ is conjugate to u. Assume that the theorem holds for all words with conjugates containing fewer syllables than u, or with the same number of syllables but of shorter length. We may consider two cases.

Case 1: $|\alpha_i| \geqslant 2$ for some i, or $|\beta_j| \geqslant 2$ for some j.

By choosing a suitable conjugate of u and inverting if necessary we may assume that $\alpha_1 \geqslant 2$ or $\beta_m \geqslant 2$. Then

$$\operatorname{tr} u = \operatorname{tr} x \operatorname{tr} x^{\alpha_1 - 1} y^{\beta_1} \dots x^{\alpha_m} y^{\beta_m} - \operatorname{tr} x^{\alpha_1 - 2} y^{\beta_1} \dots x^{\alpha_m} y^{\beta_m}$$

and

$$\operatorname{tr}(\operatorname{rev} u) = \operatorname{tr} x \operatorname{tr}(\operatorname{rev} x^{\alpha_1 - 1} y^{\beta_1} \dots x^{\alpha_m} y^{\beta_m}) - \operatorname{tr}(\operatorname{rev} x^{\alpha_1 - 2} y^{\beta_1} \dots x^{\alpha_m} y^{\beta_m})$$

in the former case; similar relations hold if $\beta_m \geqslant 2$. In either case the induction hypothesis gives immediately that $\operatorname{tr} u = \operatorname{tr}(\operatorname{rev} u)$.

Case 2: $|\alpha_i| = |\beta_i| = 1, 1 \leqslant i \leqslant m$.

If $\alpha_i = \beta_i = 1$, $1 \leqslant i \leqslant m$ then u is conjugate to $\operatorname{rev} u$. Otherwise suppose $\alpha_1 = -1$ or $\beta_m = -1$. Then

$$\operatorname{tr} x^{-1} y^{\beta_1} \dots x^{\alpha_m} y^{\beta_m} = \operatorname{tr} x \operatorname{tr} x^{\alpha_2} y^{\beta_2} \dots x^{\alpha_m} y^{\beta_1 + \beta_m} - \operatorname{tr} x y^{\beta_1} \dots x^{\alpha_m} y^{\beta_m}$$

and

$$\operatorname{tr}(\operatorname{rev} x^{-1} y^{\beta_1} \dots x^{\alpha_m} y^{\beta_m}) = \operatorname{tr} x \operatorname{tr}(\operatorname{rev} x^{\alpha_2} y^{\beta_2} \dots x^{\alpha_m} y^{\beta_1 + \beta_m}) \\ - \operatorname{tr}(\operatorname{rev} x y^{\beta_1} \dots x^{\alpha_m} y^{\beta_m}),$$

in the former case; similar relations hold if $\beta_m = -1$.

In either case, the first terms on the right-hand side are equal by the induction hypothesis, and the terms can be shown to be equal by applying the same process repeatedly until the last terms are $\operatorname{tr}(xy)^m$ and $\operatorname{tr}(\operatorname{rev}(xy)^m)$.

Example 8.2 of Horowitz (1972), showing that there is no bound on the orders of sets of non-conjugate two-variable words with the same trace, indicates that a necessary condition for two words to have the same trace must be much weaker than that of Theorem 6.3.

LEMMA 6.4 (Horowitz (1972) 2.7). *For any integer m,* $\operatorname{tr} x^m = C_{|m|}(\operatorname{tr} x)$ *where*

$C_m(z)$ is defined inductively by $C_0(z) = 2$, $C_1(z) = z$, and

$$C_m(z) = zC_{m-1}(z) - C_{m-2}, \quad m \geqslant 2.$$

LEMMA 6.5 (Horowitz (1972) Lemma 6.1). *Let u and v be cyclically reduced two-variable words of the form*

$$u = x^{\alpha_1} y^{\beta_1} x^{\alpha_2} y^{\beta_2} \dots x^{\beta_r} y^{\beta_r},$$
$$v = x^{\gamma_1} y^{\delta_1} x^{\gamma_2} y^{\delta_2} \dots x^{\gamma_s} y^{\delta_s}.$$

If $\operatorname{tr} u = \operatorname{tr} v$ *then $r = s$ and the family of syllables $\{x^{|\alpha_1|}, y^{|\beta_1|}, \dots, x^{|\alpha_r|}, y^{|\beta_r|}\}$ is the same as the family $\{x^{|\gamma_1|}, y^{|\delta_1|}, \dots, x^{|\gamma_s|}, y^{|\delta_s|}\}$.*

THEOREM 6.6. *Let*

$$u = x_{v_1}^{\alpha_1} x_{v_2}^{\alpha_2} \dots x_{v_r}^{\alpha_r}, \quad v_r \neq v_1, \quad v_i \neq v_{i+1}, \quad 1 \leqslant i < r,$$

and

$$v = x_{\mu_1}^{\beta_1} x_{\mu_2}^{\beta_2} \dots x_{\mu_s}^{\beta_s}, \quad \mu_s \neq \mu_1, \quad \mu_i \neq \mu_{i+1}, \quad 1 \leqslant i < r,$$

be cyclically reduced words in n variables. If $\operatorname{tr} u = \operatorname{tr} v$ *then $r = s$ and the family of syllables $\{x_{v_1}^{|\alpha_1|}, \dots, x_{v_r}^{|\alpha_r|}\}$ is equal to the family $\{x_{\mu_1}^{|\beta_1|}, \dots, x_{\mu_s}^{|\beta_s|}\}$.*

PROOF. By 6.4 and 6.5 the theorem is true for words in one or two variables. Suppose $n > 2$ and the result is true for all words in $n-1$ variables. Then consider the mapping θ defined by $\theta(x_i) = x_i$, $1 \leqslant i < n$, $\theta(x_n) = x_2^{-a} x_1^b x_2^a$ where $|a|$ is greater than the length of any syllable in x_2 occurring in u or v, and b is an arbitrary non-zero integer.

The images of u and v under the mapping θ must have the same trace. But under θ, all syllables in the variables $x_3, \dots, x_{n-1}$ remain distinct and unchanged, hence the family of lengths of syllables in each of these variables is the same in u as in v.

Also, under the mapping θ, all syllables in x_1 are preserved, and corresponding to each syllable of length c in x_n which occurred in the original word, a syllable in x_1 of length bc is introduced. Since the family of lengths of syllables in x_1-must be the same in the images of u and v for arbitrary b, the family of lengths of syllables in x_1 is the same in u as in v, and the family of lengths of syllables in x_n is the same in u as in v.

If we now consider the images of u and v under the mapping χ defined by $\chi(x_i) = x_i$, $1 \leqslant i < n$, $\chi(x_n) = x_1^{-a} x_2^b x_1^a$ where $|a|$ is greater than the length of any syllable in x_1 occurring in u or v, and b is an arbitrary non-zero integer, a similar argument shows that the family of lengths of syllables in x_2 is the same in u as in v.

The necessary condition given in Theorem 6.6 for $\operatorname{tr} u$ to be equal to $\operatorname{tr} v$ is much too weak. The following condition, while still not sufficient, seems plausible.

https://doi.org/10.1017/S1446788700012544 Published online by Cambridge University Press

CONJECTURE 6.7. *Let* $u = x_{v_1}^{\alpha_1} x_{v_2}^{\alpha_2} \dots x_{v_r}^{\alpha_r}$, $v_r \neq v_1$ *and* $v = x_{\mu_1}^{\beta_1} \dots x_{\mu_s}^{\beta_s}$, $\mu_s \neq \mu_1$, *be cyclically reduced words in n variables. If* $\operatorname{tr} u = \operatorname{tr} v$ *then the family of syllables of u is equal to the family of syllables of v or* v^{-1}.

There is considerable evidence favouring this hypothesis in the two-variable case. Once proved for two variables the method of proof of Theorem 6.6 would lead to the general result.

Let $u = x^{\alpha_1} y^{\beta_1} \dots x^{\alpha_r} y^{\beta_r}$ and $v = x^{\gamma_1} y^{\delta_1} \dots x^{\delta_s} y^{\delta_s}$. Then if $\operatorname{tr} u = \operatorname{tr} v$ we have

$$\left|\sum_{i=1}^{r} \alpha_i\right| = \left|\sum_{i=1}^{s} \gamma_i\right|,$$

$$\left|\sum_{i=1}^{r} \beta_i\right| = \left|\sum_{i=1}^{s} \delta_i\right| \quad \text{and} \quad \left|\sum_{i=1}^{r} (\alpha_i+\beta_i)\right| = \left|\sum_{i=1}^{s} (\gamma_i+\delta_i)\right|$$

from considering mappings which respectively

(a) map y to the identity and x to an arbitrary element of $SL(2, K)$,
(b) map x to the identity and y to an arbitrary element of $SL(2, K)$,
(c) map x and y to the same arbitrary element of $SL(2, K)$.

Now define a mapping ρ by

$$\rho(x) = \begin{Bmatrix} 1 & t \\ 0 & 1 \end{Bmatrix} \quad \text{and} \quad \rho(y) = \begin{Bmatrix} 1 & 0 \\ t & 1 \end{Bmatrix},$$

where t is an arbitrary integer. Then

$$\rho(x^{\alpha_i}) = \begin{Bmatrix} 1 & \alpha_i t \\ 0 & 1 \end{Bmatrix}, \quad \rho(y^{\beta_i}) = \begin{Bmatrix} 1 & 0 \\ \beta_i t & 1 \end{Bmatrix} \quad \text{and} \quad \rho(x^{\alpha_i} y^{\beta_i}) = \begin{Bmatrix} 1+\alpha_i \beta_i t^2 & \alpha_i t \\ \beta_i t & 1 \end{Bmatrix}.$$

THEOREM 6.8. *If* $u = x^{\alpha_1} y^{\beta_1} \dots x^{\alpha_r} y^{\beta_r}$ *then*

$$\rho(u) = \begin{bmatrix} \sum_{i=0}^{r} a_{ir} t^{2i} & \sum_{i=0}^{r-1} b_{ir} t^{2i+1} \\ \sum_{i=0}^{r-1} c_{ir} t^{2i+1} & \sum_{i=0}^{r-1} d_{ir} t^{2i} \end{bmatrix}$$

where

$$a_{0r} = 1,$$

$$a_{ir} = \sum_{\substack{1 \leqslant \mu_j \leqslant v_j \leqslant r,\ 1 \leqslant j \leqslant i \\ v_k < \mu_{k+1},\ 1 \leqslant k < i}} \left\{ \prod_{m=1}^{i} \alpha_{\mu_m} \beta_{v_m} \right\}, \quad 1 \leqslant i \leqslant r,$$

$$b_{0r} = \sum_{m=1}^{r} \alpha_m,$$

$$b_{ir} = \sum_{\substack{1 \leqslant \mu_j \leqslant v_j < r,\ 1 \leqslant j \leqslant i \\ v_k < \mu_{k+1},\ 1 \leqslant k \leqslant i}} \left\{ \alpha_{\mu_{i+1}} \prod_{m=1}^{i} \alpha_{\mu_m} \beta_{v_m} \right\}, \quad 1 \leqslant i \leqslant r-1,$$

https://doi.org/10.1017/S1446788700012544 Published online by Cambridge University Press

$$c_{0r} = \sum_{m=1}^{r} \beta_m,$$

$$c_{ir} = \sum_{\substack{1 \leqslant \mu_j < \nu_j \leqslant r,\ 1 \leqslant j \leqslant i \\ \nu_k \leqslant \mu_{k+1},\ 1 \leqslant k \leqslant i}} \left\{ \beta_{\mu_{i+1}} \prod_{m=1}^{i} \beta_{\mu_m} \alpha_{\nu_m} \right\}, \quad 1 \leqslant i \leqslant r-1,$$

$$d_{0r} = 1,$$

$$d_{ir} = \sum_{\substack{1 \leqslant \mu_j < \nu_j \leqslant r,\ 1 \leqslant j \leqslant i \\ \nu_k \leqslant \mu_{k+1},\ 1 \leqslant k < i}} \left\{ \prod_{m=1}^{i} \beta_{\mu_m} \alpha_{\nu_m} \right\}, \quad 1 \leqslant i \leqslant r-1.$$

PROOF. The proof is by induction on r. The theorem holds for $r = 1, 2$ so assume it to be true for $r = h$, say. Matrix multiplication then yields the relations (for $h > 2$)

$$\begin{aligned}
a_{0,h+1} &= 1, \\
a_{i,h+1} &= a_{ih} + \beta_{h+1} b_{i-1,h} + \alpha_{h+1} \beta_{h+1} a_{i-1,h}, \quad 1 \leqslant i \leqslant h, \\
a_{h+1,h+1} &= \alpha_{h+1} \beta_{h+1} a_{hh}, \\
b_{0,h+1} &= \alpha_{h+1} + b_{0h} \\
b_{i,h+1} &= \alpha_{h+1} a_{ih} + b_{ih}, \quad 1 \leqslant i \leqslant h-1, \\
b_{h,h+1} &= \alpha_{h+1} a_{hh}, \\
c_{0,h+1} &= \beta_{h+1} + c_{0h}, \\
c_{i,h+1} &= c_{ih} + \beta_{h+1} d_{i-1,h} + \alpha_{h+1} \beta_{h+1} c_{i-1,h}, \quad 1 \leqslant i \leqslant h-1, \\
c_{h,h+1} &= \alpha_{h+1} \beta_{h+1} c_{h-1,h}, \\
d_{0,h+1} &= 1 \\
d_{i,h+1} &= \alpha_{h+1} c_{i-1,h} + d_{ih}, \quad 1 \leqslant i \leqslant h-1, \\
d_{h,h+1} &= \alpha_{h+1} c_{h-1,h}.
\end{aligned}$$

The theorem follows immediately from these relations and the induction hypothesis.

For example, take $r = 3$. Then

$$\begin{aligned}
a_{03} &= 1, \\
a_{13} &= \alpha_1 \beta_1 + \alpha_1 \beta_2 + \alpha_1 \beta_3 + \alpha_2 \beta_2 + \alpha_2 \beta_3 + \alpha_3 \beta_3, \\
a_{23} &= \alpha_1 \beta_1 \alpha_2 \beta_2 + \alpha_1 \beta_1 \alpha_2 \beta_3 + \alpha_1 \beta_1 \alpha_3 \beta_3 + \alpha_1 \beta_2 \alpha_3 \beta_3 + \alpha_2 \beta_2 \alpha_3 \beta_3, \\
a_{33} &= \alpha_1 \beta_1 \alpha_2 \beta_2 \alpha_3 \beta_3, \\
b_{03} &= \alpha_1 + \alpha_2 + \alpha_3, \\
b_{13} &= \alpha_1 \beta_1 \alpha_2 + \alpha_1 \beta_1 \alpha_3 + \alpha_2 \beta_2 \alpha_3, \\
b_{23} &= \alpha_1 \beta_1 \alpha_2 \beta_2 \alpha_3, \\
c_{03} &= \beta_1 + \beta_2 + \beta_3, \\
c_{13} &= \beta_1 \alpha_2 \beta_2 + \beta_1 \alpha_2 \beta_3 + \beta_1 \alpha_3 \beta_3 + \beta_2 \alpha_3 \beta_3, \\
c_{23} &= \beta_1 \alpha_2 \beta_2 \alpha_3 \beta_3, \\
d_{03} &= 1, \\
d_{13} &= \beta_1 \alpha_2 + \beta_1 \alpha_3 + \beta_2 \alpha_3, \\
d_{23} &= \beta_1 \alpha_2 \beta_2 \alpha_3.
\end{aligned}$$

https://doi.org/10.1017/S1446788700012544 Published online by Cambridge University Press

Given a two-variable word u, all candidates for words with the same trace can be obtained, according to Theorem 6.6, by permuting the syllables in x and in y, and changing the signs of some of them. If the only such transformations which leave invariant the expressions a_{rr} and $a_{ir}+d_{ir}$, $0 \leqslant i \leqslant r-1$, may be regarded as permutations of the syllables of u or its inverse, then Conjecture 6.7 is proved.

In all cases, the condition that a_{rr} is invariant means that an even number of syllables must change sign, and $a_{1r}+d_{1r}$ invariant implies $(\sum_{i=1}^{r} \alpha_i)(\sum_{i=1}^{r} \beta^i)$ invariant.

References

J. Cossey, S. O. Macdonald and A. P. Street (1970), 'On the laws of certain finite groups', *J. Austral. Math. Soc.* **11,** 441–489.

R. D. Horowitz (1972), 'Characters of free groups represented in two-dimensional special linear group', *Comm. Pure Appl. Math.* **25,** 635–650.

J. B. Southcott (1974a), 'A basis for the laws of a class of simple groups', *J. Austral. Math. Soc* **17,** 500–505.

J. B. Southcott (1974b), 'Two-variable laws for a class of finite simple groups', *Bull. Austral. Math. Soc.* **10,** 85–89.

A. Whittemore (1973), 'On special linear characters of free groups of rank $n \geqslant 4$, *Proc. Amer. Math. Soc.* **40,** 383–388.

Department of Computing Science
University of Adelaide
G.P.O. Box 498
Adelaide 5001
Australia

https://doi.org/10.1017/S1446788700012544 Published online by Cambridge University Press