PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 84, Number 3, March 1982

TRACES IN 2-GENERATOR SUBGROUPS OF SL(2, C)

TROELS JORGENSEN[1]

ABSTRACT. An explicit formula for the trace polynomials of elements in 2-generator subgroups of SL(2, C) will be presented.

1. Introduction. Consider two variable elements X_1 and X_2 of SL(2, C). The traces of the elements in the group generated by X_1 and X_2 are polynomials with integer coefficients in $\tau_1 = \text{trace } X_1$, $\tau_2 = \text{trace } X_2$ and $t = \text{trace } X_1X_2$, a fact which was known to Fricke and Klein [**1**]. The interest in these polynomials has been revived by Magnus [**4**], Horowitz [**2**] and Traina [**5**]. It is the purpose to present them by an explicit formula.

2. Chebychev polynomials. Let X be an element of SL(2, C) and let m be an integer. It follows from the characteristic matrix equation $X + X^{-1} = \text{trace } X$ that

$$X^m = -\beta_{m-1} + \beta_m X, \tag{1}$$

where the functions β_n are polynomials of degree $|n| - 1$ in the variable $\tau = \text{trace } X$. These polynomials may be computed recursively using that $\beta_0 = 0$, $\beta_1 = 1$ and that

$$\tau\beta_n = \beta_{n+1} + \beta_{n-1}. \tag{2}$$

It can be verified, for instance by induction, that if n is a natural number, then

$$\beta_n(\tau) = \sum_{i=0}^{(n-1)/2} (-1)^i \binom{n-1-i}{i} \tau^{n-1-2i}. \tag{3}$$

Furthermore, one has

$$\beta_{-n} = -\beta_n. \tag{4}$$

The trace on X^n is a polynomial τ_n of degree $|n|$ in τ. It is given by

$$\tau_n = \beta_{n+1} - \beta_{n-1},$$

and may be computed recursively, using that $\tau_0 = 2$, $\tau_1 = \tau$ and

$$\tau\tau_n = \tau_{n+1} + \tau_{n-1}.$$

Later we will make use of the identity

$$\beta_{p-q} = \beta_{p-1}\beta_q - \beta_p\beta_{q-1} \tag{5}$$

and its variants

$$\beta_{p-q} = \beta_p\beta_{q+1} - \beta_{p+1}\beta_q, \tag{6}$$

Received by the editors March 16, 1981.
1980 *Mathematics Subject Classification.* Primary 20H05; Secondary 30F40.
[1]Supported in part by the National Science Foundation and by the Sloan Foundation.

0002-9939/81/0000-1017/$02.25

(7) $$\beta_{p-q+1} = \beta_p\beta_q - \beta_{p+1}\beta_{q-1},$$

(8) $$\beta_{p-q-1} = \beta_{p-1}\beta_{q+1} - \beta_p\beta_q.$$

It may be observed that if one writes $\tau = \lambda + \lambda^{-1}$, then

$$\beta_n = (\lambda^n - \lambda^{-n})(\lambda - \lambda^{-1})^{-1} \quad \text{and} \quad \tau_n = \lambda^n + \lambda^{-n}.$$

3. Notation. Given X_1 and X_2, we put

$$(X_i, \tau_i) = \begin{cases} (X_1, \tau_1) & \text{if } i \text{ is odd,} \\ (X_2, \tau_2) & \text{if } i \text{ is even} \end{cases}$$

and

$$t_n = \operatorname{trace}(X_1X_2)^n.$$

Furthermore, if $\mathbf{m}$ is a vector with $2r$ integer coordinates m_i, we define

$$B(\mathbf{m}) = \prod_{i=1}^{2r} \beta_{m_i}(\tau_i)$$

and

$$T(\mathbf{m}) = \operatorname{trace} \prod_{i=1}^{2r} X_i^{m_i}.$$

4. Motivation. With use of (1), we obtain

$$T(\mathbf{m}) = \operatorname{trace} \prod_{i=1}^{2r} \left(-\beta_{m_i-1}(\tau_i) + \beta_{m_i}(\tau_i)X_i\right) = \sum_{\boldsymbol{\delta}} (-1)^{\Sigma\delta_i} B(\mathbf{m} + \boldsymbol{\delta} - \mathbf{1})T(\boldsymbol{\delta}),$$

where $\mathbf{1}$ denotes the $2r$-vector all of whose coordinates are 1, and the summation is extended over all vectors $\boldsymbol{\delta}$ with $2r$ coordinates $\delta_i \in \{0, 1\}$. It is easy to see that the degree of $T(\boldsymbol{\delta})$ is at most

$$\sum_{i \text{ odd}} \delta_i \text{ in } \tau_1 \quad \text{and} \quad \sum_{i \text{ even}} \delta_i \text{ in } \tau_2.$$

This fact allows us to obtain a formula of the type

(∗) $$T(\mathbf{m}) = \sum_{\boldsymbol{\epsilon}} f_{\boldsymbol{\epsilon}} B(\mathbf{m} + \boldsymbol{\epsilon}),$$

where $\boldsymbol{\epsilon}$ runs through all vectors with $2r$ coordinates belonging to $\{-1, 0, 1\}$, and the coefficients $f_{\boldsymbol{\epsilon}}$ are polynomials in $t = \operatorname{trace} X_1X_2$. Loosely speaking, the coefficient $B(\mathbf{m} + \boldsymbol{\delta} - \mathbf{1})$ to $T(\boldsymbol{\delta})$ contains enough factors β_{m_i} to "absorb" the powers of τ_i using (2), that is $\tau_i\beta_{m_i} = \beta_{m_i+1} + \beta_{m_i-1}$. Since this procedure can be carried out in different ways, the polynomials $f_{\boldsymbol{\epsilon}}$ are not unique.

5. Symmetry. A formula of that type (∗) is said to be symmetric if $f_{-\boldsymbol{\epsilon}} = f_{\boldsymbol{\epsilon}}$ for all $\boldsymbol{\epsilon}$. It is always possible to symmetrize. Namely, if

$$T(\mathbf{m}) = \sum_{\boldsymbol{\epsilon}} f_{\boldsymbol{\epsilon}} B(\mathbf{m} + \boldsymbol{\epsilon}),$$

then also

$$T(\mathbf{m}) = \frac{1}{2} \sum_{\boldsymbol{\epsilon}} (f_{\epsilon} + f_{-\epsilon}) B(\mathbf{m} + \boldsymbol{\epsilon}).$$

This is so because of the identities

$$T(\mathbf{m}) = T(-\mathbf{m}) \quad \text{and} \quad B(-\mathbf{m} + \boldsymbol{\epsilon}) = B(\mathbf{m} - \boldsymbol{\epsilon}).$$

The latter follows from (4). To prove the former, one may use that $X_1 X_2 - X_2 X_1$ generically is an invertible matrix, which conjugates X_1 into X_1^{-1} and X_2 into X_2^{-1}, hence $\prod X_i^{m_i}$ into $\prod X_i^{-m_i}$ (for instance, see [3]). Since the trace is invariant under conjugation, the identity follows.

6. The degree. For a vector $\boldsymbol{\epsilon} = (\varepsilon_1, \ldots, \varepsilon_{2r})$, $\varepsilon_i \in \{-1, 0, 1\}$, we denote by $|\boldsymbol{\epsilon}|$ the vector with coordinates $|\varepsilon_i|$. The degree $d(\boldsymbol{\epsilon})$ of $\boldsymbol{\epsilon}$ is defined as the degree of $T(\mathbf{1} - |\boldsymbol{\epsilon}|)$ in the variable $t = \operatorname{trace} X_1 X_2$. Choosing X_1 and X_2 with trace 0, one may verify the following recipe for computing the degree: Think of $\boldsymbol{\epsilon}$ as a cycle and "cut out" each subinterval consisting of a maximal and even number of consecutive nonzero coordinates. For each of the remaining intervals, form the alternating sum of the number of consecutive zeroes. Half the sum of the absolute values of these sums is $d(\boldsymbol{\epsilon})$.

EXAMPLE, $r = 10$.

$$\boldsymbol{\epsilon} = (+, -, 0, 0, 0, 0, +, 0, -, +, 0, -, +, 0, 0, -, 0, 0, 0, 0),$$

$$d(\boldsymbol{\epsilon}) = \tfrac{1}{2}(|4 - 1| + |1| + |2 - 4|) = 3.$$

7. Alternation. A vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_{2r})$, $\mu_i \in \{-1, 0, 1\}$ is said to be alternating if its nonzero coordinates appear with alternating signs from left to right and their number is even. A formula of the type $(*)$ is called alternating if $f_{\epsilon} = 0$ for all $\boldsymbol{\epsilon}$ which are not alternating.

8. The formula. There is a unique symmetric, alternating formula of the type $(*)$. It is

$$T(\mathbf{m}) = \tfrac{1}{2} t_r B(\mathbf{m}) + \tfrac{1}{2} \sum_{\boldsymbol{\mu}} (-1)^{r - d(\boldsymbol{\mu})} t_{d(\boldsymbol{\mu})} B(\mathbf{m} + \boldsymbol{\mu}),$$

where $\boldsymbol{\mu}$ runs through the alternating $2r$-vectors. The proof of this observation occupies the next two sections.

9. Uniqueness. Suppose that there is such a formula, say

$$T(\mathbf{m}) = \sum_{\boldsymbol{\mu}} g_{\mu} B(\mathbf{m} + \boldsymbol{\mu}).$$

Let us choose X_1 and X_2 of order 4. This means that $\tau_1 = \tau_2 = 0$ and hence, by (2), that

$$\beta_n = \begin{cases} -1 & \text{if } n = 4k - 1, \\ 0 & \text{if } n \text{ is even}, \\ 1 & \text{if } n = 4k + 1. \end{cases} \tag{9}$$

In particular, it follows from (1) that

(10) $$X_1^2 = X_2^2 = -1.$$

For a given alternating vector $\boldsymbol{\theta}$ we obtain with use of (9)

$$B(\mathbf{1} - |\boldsymbol{\theta}| + \boldsymbol{\mu}) = g_{\boldsymbol{\mu}}\begin{cases} (-1)^{(1/2)\Sigma|\mu_i|} & \text{if } \mu = \pm\theta, \\ 0 & \text{otherwise,} \end{cases}$$

and, consequently,

(11) $$T(\mathbf{1} - |\boldsymbol{\theta}|) = \begin{cases} g_{\mathbf{0}} & \text{if } \boldsymbol{\theta} = \mathbf{0}, \\ g_{\boldsymbol{\theta}} 2(-1)^{(1/2)\Sigma|\theta_i|} & \text{otherwise.} \end{cases}$$

For the left side, we find with use of (10) and a little effort:

(12) $$T(\mathbf{1} - |\boldsymbol{\theta}|) = t_{d(\boldsymbol{\theta})}(-1)^{r-(1/2)\Sigma|\theta_i| - d(\boldsymbol{\theta})}.$$

Thus, comparing (11) and (12),

$$g_{\boldsymbol{\theta}} = \begin{cases} t_r & \text{if } \boldsymbol{\theta} = \mathbf{0}, \\ \frac{1}{2}(-1)^{r-d(\boldsymbol{\theta})} t_{d(\boldsymbol{\theta})} & \text{otherwise,} \end{cases}$$

which proves the uniqueness.

10. Existence. An elementary computation, making use of (1) and (2), yields

(13) $$\operatorname{trace} X_1^p X_2^q = \beta_p \beta_q t - \beta_{p-1}\beta_{q+1} - \beta_{p+1}\beta_{q-1}.$$

For convenience, we have suppressed the variables τ_1 and τ_2. This shows that "The Formula" is valid when $r = 1$. The general case is proved by induction, based on the well-known identity

(14) $$\operatorname{trace} AB = \operatorname{trace} A \operatorname{trace} B - \operatorname{trace} AB^{-1}.$$

Given a "word"

$$W = \prod_{i=1}^{2n+2} X_i^{m_i},$$

we split it into a product of two elements

$$W = AB,$$

where

$$A = \prod_{i=1}^{2n} X_i^{m_i} \quad \text{and} \quad B = X_1^{m_{2n+1}} X_2^{m_{2n+2}}.$$

Assuming that "The Formula" has been proved for $r = n$, it may be applied to A and to B, as well as to AB^{-1}, since

$$\operatorname{trace} AB^{-1} = \operatorname{trace}\left(X_1^{m_1 - m_{2n+1}} \prod_{i=2}^{2n-1} X_i^{m_i} X_2^{m_{2n} - m_{2n+2}} \right).$$

It follows that the product of the traces of A and B is expressible as a sum over certain alternating $(2n + 2)$-vectors $\boldsymbol{\mu}$, plus a sum over all $(2n + 2)$-vectors $\boldsymbol{\nu} = (\mathbf{u}, \mathbf{v})$, whose first $2n$ coordinates can be any nonzero alternating vector $\mathbf{u}$ and whose last two coordinates is the unique alternating 2-vector $\mathbf{v}$ for which $(\mathbf{u}, \mathbf{v})$ is not alternating:

$$\text{(15)} \qquad \operatorname{trace} A \operatorname{trace} B = \sum_{\boldsymbol{\mu}} f_{\boldsymbol{\mu}} B_{n+1}(\mathbf{m} + \boldsymbol{\mu}) + \sum_{\boldsymbol{\nu}} f_{\boldsymbol{\nu}} B_{n+1}(\mathbf{m}, \boldsymbol{\nu}).$$

As usual, the coefficients $f_{\boldsymbol{\mu}}$ and $f_{\boldsymbol{\nu}}$ are polynomials in t, independent of τ_1 and τ_2. Specifically, in the second sum we find that

$$f_{\boldsymbol{\nu}} = f_{(\mathbf{u}, \boldsymbol{\nu})} = \tfrac{1}{2}(-1)^{n+1-d(\mathbf{u})} t_{d(\mathbf{u})}.$$

The expression obtained for the trace of AB^{-1} appears at first as a sum over the alternating $2n$-vectors. It may be rewritten to suit our purposes as a sum over $(2n + 2)$-vectors with use of (5), (6), (7) and (8), applied to the first and the last of the factors of the B-products. Letting $\mathbf{M}$ denote the $2n$-vector, whose ith coordinate is $m_1 - m_{2n+1}$ if $i = 1$, m_i if $i = 2, \ldots, 2n - 1$ and $m_{2n} - m_{2n+2}$ if $i = 2n$, we obtain identities of the form

$$B_n(\mathbf{M} + \mathbf{u}) = B_{n+1}(\mathbf{m} + \boldsymbol{\mu}_1) - B_{n+1}(\mathbf{m} + \boldsymbol{\mu}_2) - B_{n+1}(\mathbf{m} + \boldsymbol{\mu}_3) + B_{n+1}(\mathbf{m} + \boldsymbol{\mu}_4),$$

where $\mathbf{u}$ is any alternating $2n$-vector and $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$ and $\boldsymbol{\mu}_4$ are alternating $(2n + 2)$-vectors, and $\boldsymbol{\mu}_3$ is the $(2n + 2)$-vector whose first $2n$ coordinates are equal to $\mathbf{u}$ and whose last two coordinates is the unique alternating 2-vector making $\boldsymbol{\mu}_3$ nonalternating, except when $\mathbf{u} = \mathbf{0}$, in which case also $\boldsymbol{\mu}_3$ is alternating. Keeping track of the coefficients $f_{\mathbf{u}}$, we obtain

$$\text{(16)} \qquad \operatorname{trace} AB^{-1} = \sum_{\boldsymbol{\mu}} g_{\boldsymbol{\mu}} B_{n+1}(\mathbf{m} + \boldsymbol{\mu}) + \sum_{\boldsymbol{\nu}} f_{\boldsymbol{\nu}} B_{n+1}(\mathbf{m} + \boldsymbol{\nu}),$$

where the first sum extends over alternating $(2n + 2)$-vectors, and the second sum, term by term, is equal to the second sum on the right side in (15). Subtracting (16) from (15), we obtain, because of (14), an alternating formula for the trace of $W = AB$. When symmetrized, the unique formula of §8 must be the result. Thus, "The Formula" holds for $r = n + 1$ and therefore it is valid for all r.

References

1. R. Fricke and F. Klein, *Vorlesungen über die Theorie der automorphen Functionen*, Teubner, Leipzig, 1897.

2. R. Horowitz, *Characters of free groups represented in the two-dimensional linear group*, Comm. Pure Appl. Math. **25** (1972), 635–649.

3. T. Jorgensen, *Closed geodesics on Riemann surfaces*, Proc. Amer. Math. Soc. **72** (1978), 140–142.

4. W. Magnus, *Two generator subgroups of* PSL(2, C), Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II **7** (1975), 1–14.

5. C. R. Traina, *Trace polynomials for two generator subgroups of* SL(2, C), Proc. Amer. Math. Soc. **79** (1980), 369–372.

Department of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455

Department of Mathematics, Columbia University, New York, New York 10027 (Current address)