Can. J. Math., Vol. XXVI, No. 6, 1974, pp. 1412-1417

PRODUCTS OF TRANSVECTIONS

B. B. PHADKE

1. Introduction. This paper is concerned with the presentation of certain elements of the group $SL(n, K)$ as products of a minimal number of transvections. To explain the terminology, let V be an n-dimensional left vector space over a (not necessarily commutative) field K. The group of all nonsingular linear transformations of V onto V (i.e. the group of all collineations of V) is the group $GL(n, K)$. This group is generated by collineations leaving a hyperplane pointwise fixed. When $n = 2$ these collineations are called axial collineations and the invariant hyperplane (line) is then called an axis. Among the collineations leaving a hyperplane pointwise fixed are "transvections" which move every vector by a vector lying in the invariant hyperplane. Thus a transvection T whose invariant hyperplane is $\mathscr{U}$ has the form $Tx = x + uf(x)$ where $u \in \mathscr{U}$ and f is a linear map of V into K such that $f^{-1}(0) = \mathscr{U}$. We say that the transvection T is parallel to $\mathscr{U}$. The group generated by all the transvections is the group $SL(n, K)$. Dieudonné [**4**, p. 152] states that every element A of the group $SL(n, K)$ is a product of at most $(r + 1)$ transvections where r is the rank of $A - I$, I being the identity transformation (i.e. r is the dimension of the path of A). This is false if the ground field K is not commutative. In this paper we show that certain elements of $SL(n, K)$ are products of $(r + 2)$ but not $(r + 1)$ transvections.

The reason why complications arise when K is not commutative can be explained as follows. Let D be a collineation which leaves a hyperplane $\mathscr{U}$ pointwise fixed. If D is neither the identity nor a transvection then there exists a unique 1-dimensional subspace (i.e. a line) say $W = Kv$ such that $W \cap \mathscr{U} = \{0\}$ and such that D leaves W fixed (as a line, not pointwise). Such a collineation D can be given by $Du = u$ for all u in $\mathscr{U}$ and $Dv = \alpha v$, for a certain v not in $\mathscr{U}$. Such a collineation is called a dilatation. If A is an element of $GL(n, K)$ and r is the dimension of the path of A then Dieudonné [**4**, Theorem 1, p. 149] proves that A can be written as a product of at most r transvections and a single transvection or a single dilatation. Using Dieudonné's theory of determinants over a noncommutative field (see [**2**]), we find that $SL(n, K)$ is the group of all collineations with determinant unity, see [**1**, Theorem 4.6, p. 163]. When K is commutative there are no proper dilatations in $SL(n, K)$. In other words, if K is commutative the only collineation in $SL(n, K)$ which leaves a hyperplane pointwise fixed and which is not a proper transvection is the identity. On the other hand when K is not com-

Received May 22, 1973 and in revised form, July 23, 1973.

https://doi.org/10.4153/CJM-1974-135-0 Published online by Cambridge University Press

mutative the dilatation D ($Du = u$ for all u in $\mathcal{U}$ and $Dv = \alpha v$ for a v not in V) belongs to $SL(n, K)$ whenever α belongs to the commutator subgroup of $K^* = K - \{0\}$ (see [**1**, Theorem 4.2, p. 155]). Among such dilatations are those belonging to commutators i.e. those for which $\alpha = \mu\nu\mu^{-1}\nu^{-1}$ for some μ, ν in K. This explains why complications enter when K is not commutative. It also shows that the key to expressing elements of $SL(n, K)$ as products of a minimal number of transvections lies in expressing the dilatations in $SL(n, K)$ as products of a minimal number of transvections. In this paper we prove the following theorem. We assume that $n \geqq 2$.

THEOREM. *A dilatation belonging to a commutator can always be expressed as a product of three transvections. Furthermore, given a hyperplane which does not contain the invariant line of the dilatation and which is distinct from the pointwise invariant hyperplane of the dilatation, we can choose one of the three transvections to be parallel to this given hyperplane. A non-identity dilatation is never a product of two transvections.*

2. To make the writing of the proof easier we will restrict ourselves to the 2-dimensional case in this and the next two sections. The transition to the n-dimensional case is then straightforward and is made in section 5. Greek letters and the letters a, b, c will denote elements of the field K. We identify a collineation with the matrix representing it. Thus taking the axis as a line spanned by a basis vector, a transvection can be represented by

$$T(\alpha) = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \quad \text{or} \quad T^*(\beta) = \begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix}.$$

For a dilatation we take the axis and the invariant line to be spanned by basis vectors. The matrix of a dilatation then takes the form

$$D(\mu) = \begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix}.$$

Changing the basis amounts to changing the matrix A to $B^{-1}AB$ where B is in $GL(n, K)$.

3. The matrix of a transvection. When the axis of a transvection is spanned by the transpose of the vector $(1, 0)$ the matrix is $T^*(\mu)$ for some μ while the matrix of a transvection whose axis is spanned by the transpose of $(0, 1)$ is $T(\lambda)$ for some λ. Since $T^{-1}(\lambda) = T(-\lambda)$, the transformation $T(\lambda, \mu) = T(\lambda)T^*(\mu)T(-\lambda)$ is also a transvection. On computation we find that

$$\begin{aligned} T(\lambda, \mu) = T(\lambda)T^*(\mu)T(-\lambda) &= \begin{bmatrix} 1 & 0 \\ \lambda & 1 \end{bmatrix}\begin{bmatrix} 1 & \mu \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -\lambda & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 - \mu\lambda & \mu \\ -\lambda\mu\lambda & 1 + \lambda\mu \end{bmatrix}. \end{aligned}$$

https://doi.org/10.4153/CJM-1974-135-0 Published online by Cambridge University Press

Hence the transvection $T(\lambda, \mu)$ has its axis spanned by $(1, \lambda)$ and it takes the vector $u = (0, 1)$ into $u + (1, \lambda)\mu$. We have thus proved the following lemma.

LEMMA. *The matrix of a transvection can always be expressed as either* $T(\lambda)$ *or* $T(\lambda, \mu)$.

It is easy to verify the following relations.
(3.1) $T^{-1}(\lambda) = T(-\lambda)$, $T^{*-1}(\mu) = T^*(-\mu)$
(3.2) $T(\lambda, \mu)T(\lambda, \nu) = T(\lambda, \mu + \nu)$
(3.3) $T^{-1}(\lambda, \mu) = T(\lambda, -\mu)$.

4. Proof of the theorem. Before proving the theorem we explain the method of proof. First note that $D(\sigma)$ and $D(\tau)$ are conjugate in $SL(2, K)$ when σ and τ are conjugate in K^* (see [**3**, p. 37]). Hence $D(\nu\mu^{-1}\nu^{-1}) = RD(\mu^{-1})R^{-1}$ where $R \in SL(2, K)$, i.e., R is a product of transvections. Since $(D(\mu))^{-1} = D(\mu^{-1})$ we have

$$D(\mu\nu\mu^{-1}\nu^{-1}) = D(\mu)D(\nu\mu^{-1}\nu^{-1}) = D(\mu)RD(\mu^{-1})R^{-1} = R^*R^{-1}$$

where $R^* = D(\mu)RD(\mu^{-1})$ is also a product of the same number of transvections as is R. Our method of expressing $D(\mu\nu\mu^{-1}\nu^{-1})$ as a product of 3 transvections consists in making a judicious choice for transvections which are factors of R so that several cancellations and contractions are possible in the product $R^*R^{-1} = D(\mu\nu\mu^{-1}\nu^{-1})$.

4.1. *Construction of* R. For any nonzero λ in K define

$$R(\lambda) = \begin{bmatrix} \lambda\nu^{-1}\lambda^{-1} & 0 \\ 0 & \nu \end{bmatrix}.$$

Then it can be verified that

$$D(\nu\mu^{-1}\nu^{-1}) = R(\lambda)D(\mu^{-1})(R(\lambda))^{-1}.$$

On the other hand we have

$$\begin{bmatrix} \lambda\nu^{-1}\lambda^{-1} & 0 \\ 0 & \nu \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -(1-\nu)^2\lambda^{-1} & 1 \end{bmatrix}\begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2-\lambda\nu\lambda^{-1} & -\lambda\nu \\ (1-\nu)^2\nu^{-1}\lambda^{-1} & \nu \end{bmatrix}.$$

The last equation can be expressed as

$$R(\lambda) = T(-\alpha)T^*(\lambda)T(\lambda^*, \mu^*)$$

where $\alpha = (1 - \nu)^2\lambda^{-1}$, $\lambda^* = (\nu^{-1} - 1)\lambda^{-1}$, $\mu^* = -\lambda\nu$.

4.2. LEMMA. *The following identities hold.*
(4.21) $T(a)D(\sigma)T(b) = T(c)D(\sigma)T(-c)$, *where* $c = (1 - \sigma)^{-1}(a + \sigma b)$.
(4.22) $T^*(a)D(\sigma)T^*(b) = T^*(c)D(\sigma)T^*(-c)$, *where* $c = (b + a\sigma)(1 - \sigma)^{-1}$.
(4.23) $D(\sigma)T(a)D(\sigma^{-1}) = T(\sigma a)$.
(4.24) $D(\sigma)T^*(a)D(\sigma^{-1}) = T^*(a\sigma^{-1})$.

The proofs can be obtained by straightforward computation.

4.3. We write $\sigma = \mu\nu\mu^{-1}\nu^{-1}$. α, λ^*, μ^* are as in § 4.1. Put

$$T_1 = D(\mu)T(\lambda^*, \mu^*)D(\mu^{-1}),\ T_2 = T^{-1}(\lambda^*, \mu^*).$$

Then using § 4.1 and § 4.2 we have

$$\begin{aligned} D(\sigma) &= D(\mu\nu\mu^{-1}\nu^{-1}) \\ &= D(\mu)\cdot R(\lambda)\cdot D(\mu^{-1})\cdot (R(\lambda))^{-1} \\ &= (D(\mu)T(-\alpha)T^*(\lambda)T(\lambda^*, \mu^*)\cdot D(\mu^{-1}))\cdot T^{-1}(\lambda^*, \mu^*)\cdot \\ &\qquad T^{*-1}(\lambda)T^{-1}(-\alpha) \\ &= T(-\mu\alpha)T^*(\lambda\mu^{-1})T_1T_2T^*(-\lambda)T(\alpha). \end{aligned}$$

In the last equation we used

$$T(-\mu\alpha) = D(\mu)T(-\alpha)D(\mu^{-1})$$

and

$$T^*(\lambda\mu^{-1}) = D(\mu)T^*(\lambda)D(\mu^{-1}).$$

Thus we have

$$T(\mu\alpha)D(\sigma)T(-\alpha) = T^*(\lambda\mu^{-1})T_1T_2T^*(-\lambda).$$

Using § 4.21, $T(\mu\alpha)D(\sigma)T(-\alpha) = T(\epsilon)D(\sigma)T(-\epsilon)$ where

$$\epsilon = (1-\sigma)^{-1}(\mu\alpha - \sigma\alpha) = (1-\sigma)^{-1}(\mu-\sigma)\alpha.$$

Hence

$$\begin{aligned} D(\sigma) &= T(-\epsilon)T^*(\lambda\mu^{-1})T_1T_2T^*(-\lambda)T(\epsilon) \\ &= T(-\epsilon)T^*(\lambda\mu^{-1})T_1T_2T^*(-\lambda\mu^{-1})\cdot T(\epsilon)\cdot \\ &\qquad T(-\epsilon)T^*(\lambda\mu^{-1})T^*(-\lambda)T(\epsilon) \\ &= T_3T_4T_5, \end{aligned}$$

where

$$T_3 = T(-\epsilon)T^*(\lambda\mu^{-1})T_1T^*(-\lambda\mu^{-1})T(\epsilon)$$
$$T_4 = T(-\epsilon)T^*(\lambda\mu^{-1})T_2T^*(-\lambda\mu^{-1})T(\epsilon)$$

and

$$T_5 = T(-\epsilon)T^*(\lambda\mu^{-1})T^*(\lambda)T(\epsilon) = T(-\epsilon)T^*(\lambda\mu^{-1}+\lambda)T(\epsilon).$$

Thus we have proved that $D(\mu\nu\mu^{-1}\nu^{-1})$ is always a product of 3 transvections.

4.4. We now prove that we can choose λ in such a way that T_5 is parallel to any given direction other than the directions spanned by $(1, 0)$ and $(0, 1)$. To show this write $\gamma = \lambda\mu^{-1} + \lambda$ and observe that

$$T_5 = T(-\epsilon)T^*(\gamma)T(\epsilon) = T(-\epsilon, \gamma) \text{ by § 3}.$$

Hence T_5 is parallel to the direction spanned by $(1, -\epsilon)$. Since

$$\begin{aligned} \epsilon &= (1-\sigma)^{-1}(\mu-\sigma)\alpha \\ &= (1-\sigma)^{-1}(\mu-\sigma)(1-\nu)^2\lambda^{-1}, \end{aligned}$$

https://doi.org/10.4153/CJM-1974-135-0 Published online by Cambridge University Press

λ can be chosen in such a way that ϵ is equal to any given nonzero element of K. This implies that T_5 can be made to be parallel to any given direction other than those spanned by $(1, 0)$ and $(0, 1)$.

4.5. Finally we prove that the dilatation $D(\mu) = D$ can never be a product of 2 transvections. Write $u = (1, 0)$ so that $Du = u$. Hence if T_1 and T_2 are 2 transvections such that $D = T_1T_2$ then $Du = u = T_1T_2u$ so that $T_2u = T_1^{-1}u$. Therefore $v = T_2u - u = T_1^{-1}u - u$. Now T_1 and T_2 must have distinct axes, $\mathscr{U}_1$ and $\mathscr{U}_2$ say, for if $\mathscr{U}_1 = \mathscr{U}_2$ then $D = T_1T_2$ would be a transvection also. But when $\mathscr{U}_1 \neq \mathscr{U}_2$ we have, since $\mathscr{U}_1$ and $\mathscr{U}_2$ are lines through the origin, $\mathscr{U}_1 \cap \mathscr{U}_2 = 0$ so that $v \in \mathscr{U}_1 \cap \mathscr{U}_2$ implies $v = 0$. This means $T_2u = u$ and $T_1^{-1}u = u$, i.e., $T_1u = u$; which implies that u spans the axes of both T_1 and T_2 which is a contradiction to $\mathscr{U}_1 \neq \mathscr{U}_2$.

The results of § 4.3, § 4.4 and § 4.5 completely prove the theorem in two dimensions.

5. The transition to n-dimensions. All the matrix calculations in § 2, § 3 and § 4 can be carried over to the n-dimensional case by replacing a 2×2 matrix M by the $n \times n$ matrix $\tilde{M}$ given by

$$\tilde{M} = \begin{bmatrix} I_{n-2} & 0 \\ 0 & M \end{bmatrix}$$

where I_{n-2} is the identity matrix of order $(n - 2) \times (n - 2)$ and all other entries are zeros. Hence the calculations of § 4 show that every dilatation can be expressed as a product of three transvections. To get the complete result of the theorem let D be a dilatation with invariant hyperplane $\mathscr{U}$ and invariant line spanned by v. Let W be any other hyperplane such that $v \notin W$. Then $X = \mathscr{U} \cap W$ is a subspace of dimension $(n - 2)$. We can choose a basis $\{e_1, \ldots, e_n\}$ in such a way that $e_n = v$, the set $\{e_1, \ldots, e_{n-1}\}$ spans $\mathscr{U}$ and $\{e_1, \ldots, e_{n-2}, w\}$ spans W, where $w \neq e_{n-1}$, $w \neq v = e_n$. Then the results of § 4 show that D can be expressed as a product of 3 transvections each of which keeps X pointwise fixed and one of which fixes w. Thus we have proved that one of the transvections can be chosen to be parallel to any given hyperplane which is distinct from the invariant hyperplane of the dilatation and which does not contain the invariant line of the dilatation.

It remains to show that a dilatation cannot be expressed as a product of two transvections. To prove this let $\mathscr{U}$, $\mathscr{U}_1$, $\mathscr{U}_2$ be respectively the pointwise invariant hyperplanes of a dilatation D and transvections T_1 and T_2 and assume that $D = T_1T_2$. Then for every $u \in \mathscr{U} \cap \mathscr{U}_2$ we have $u = Du = T_1T_2u = T_1u$, showing that u is kept fixed by T_1 also. Thus $\mathscr{U}_1 \supset \mathscr{U} \cap \mathscr{U}_2$ so that $\mathscr{U} \cap \mathscr{U}_1 \cap \mathscr{U}_2 = W$ is a subspace of dimension at least $(n - 2)$. Hence we can choose a basis $\{e_1, \ldots, e_n\}$ of V such that the set $\{e_1, \ldots, e_{n-1}\}$ spans $\mathscr{U}$, e_n spans the invariant line of D and $e_1, \ldots, e_{n-2}$ are in $\mathscr{U}_1 \cap \mathscr{U}_2$. Thus we are back in the case of § 4.5 and from the proof given there we conclude that a proper dilatation can never be expressed as a product of two transvections. This completes the proof of the theorem in n-dimensions.

I am grateful to Professor Peter Scherk for helpful discussions.

References

1. E. Artin, *Geometric algebra* (Interscience Publishers, New York, 1957).

2. J. Dieudonné, *Les determinants sur un corps non commutatif*, Bull. Soc. Math. France *71* (1943), 27–45.

3. ——— *La géométrie des groupes classiques*, 2nd ed., Ergebnisse der Math. Neue Folge, Band 5 (Springer-Verlag, Berlin 1963).

4. ——— *Sur les générateurs des groupes classiques*, Summa Brasiliensis Mathematicae *3* (1955), 149–180.

The Flinders University of South Australia,
Bedford Park, South Australia