PACIFIC JOURNAL OF MATHEMATICS
Vol 104, No 1, 1983

LINEAR TRANSFORMATIONS THAT PRESERVE THE NILPOTENT MATRICES

PETER BOTTA, STEPHEN PIERCE AND WILLIAM WATKINS

Let sl_n be the algebra of $n \times n$ matrices with zero trace and entries in a field with at least n elements. Let $\mathfrak{N}$ be the set of nilpotent matrices. The main result in this paper is that the group of nonsingular linear transformations L on sl_n such that $L(\mathfrak{N}) = \mathfrak{N}$ is generated by the inner automorphisms: $X \to S^{-1}XS$; the maps: $X \to aX$, for $a \neq 0$; and the map: $X \to X^t$ that sends a matrix X to its transpose.

Introduction. Let M_n be the algebra of $n \times n$ matrices over a field K and let S be an algebraic set in M_n. There are a number of theorems characterizing the linear maps L on M_n that preserve S, i.e. $L(S) \subseteq S$. For example there are results for $\{X\colon \det X = 0\}$ by Dieudonné [**4**], $\{X\colon \operatorname{rank} X \leq 1\}$ by Jacob [**8**] and Marcus and Moyls [**10**], the orthogonal group by Pierce and Botta [**2**] and other linear groups by Dixon [**5**]. In every instance the transformations L that preserve these various algebraic sets have one of these two forms:

$$(1) \qquad L(X) = PXQ, \qquad \text{for all } X$$

or

$$(2) \qquad L(X) = PX^tQ, \qquad \text{for all } X$$

where P and Q are in M_n. There are conditions on P and Q which depend on the algebraic set S. For example if $S = \{X\colon \det X = 0\}$ and L is nonsingular then P and Q are nonsingular; if S is the orthogonal group then $PQ = I$ and P must be a scalar multiple of a matrix in the orthogonal group over the algebraic closure of K. For a good survey of further results of this type see Marcus [**9**].

In this paper we characterize the nonsingular linear transformations L that preserve the set $\mathfrak{N}$ of nilpotent matrices. Since the linear span of $\mathfrak{N}$ is the space sl_n of matrices with trace zero, we may as well assume that L is a transformation on sl_n. (In order to see that $\mathfrak{N}$ spans sl_n, let E_{ij} be the matrix whose only nonzero entry is a 1 in position (i, j). The nilpotent matrices E_{ij} and $E_{ii} + E_{ij} - E_{ji} - E_{jj}$ for $i \neq j$ span sl_n.)

Actually we characterize all nonsingular semilinear mappings that preserve nilpotence. The main theorem can be extended to matrices with entries from an integral domain. The extension follows from a modification of a result of Chevalley [**3**, p. 104, Théorème 3].

THEOREM. *Let $n \geq 3$, K be a field with at least n elements and suppose that L is a nonsingular linear transformation on sl_n such that $L(\mathfrak{N}) \subseteq \mathfrak{N}$. Then L either has form* (1) *or* (2), *where PQ is a non-zero scalar matrix.*

Without the assumption that L is nonsingular the theorem is false. Any map whose image is contained in the algebra $\mathfrak{U}$ of the strictly upper triangular matrices preserves nilpotence. The proof of the theorem depends on a result of Gerstenhaber about maximal spaces of nilpotent matrices. We also use some elementary algebraic geometry and the fundamental theorem of projective geometry [**1**, p. 88, Theorem 2.26].

LEMMA 1 (Gerstenhaber [**6**]). *Suppose K has at least n elements and $\mathfrak{N}$ is a space of nilpotent matrices. Then $\dim \mathfrak{N} \leq n(n-1)/2$. If $\dim \mathfrak{N} = n(n-1)/2$, then there exists a non-singular matrix S such that $\mathfrak{N} = S^{-1}\mathfrak{U}S$, where $\mathfrak{U}$ is the algebra of strictly upper triangular matrices. Moreover, any matrix of nilindex n is contained in exactly one maximal nilpotent algebra.*

Tangent Spaces. Let $K[X] = K[X_{11}, \ldots, X_{nn}]$ be the ring of polynomials in n^2 variables with coefficients in K. For $r = 1, 2, \ldots, n$, let $E_r(X) \in K[X]$ be the rth elementary symmetric function of the matrix $X = (X_{ij})$, i.e. $E_r(X)$ is the sum of all principal $r \times r$ subdeterminants of X. We let J be the ideal in $K[X]$ generated by $E_1(X), \ldots, E_n(X)$ and $\mathrm{rad}\, J = \{F \in K[X]\colon F^k \in J$ for some positive integer $k\}$. Clearly we have $\mathfrak{N} = \{A \in M_n\colon F(A) = 0$ for all $F \in J\}$. If $A \in \mathfrak{N}$ then

$$\tan(J, A) = \left\{B \in M_n\colon \frac{dF}{dt}(A + tB)\Big|_{t=0} = 0 \quad \text{for all } F \in J\right\}$$

and

$$\tan(\mathrm{rad}\, J, A) = \left\{B \in M_n\colon \frac{dF}{dt}(A + tB)\Big|_{t=0} = 0 \quad \text{for all } F \in \mathrm{rad}\, J\right\}.$$

Both of these are vector spaces and the second is the usual tangent space at the point A of the algebraic set $\mathfrak{N}$. Further, the second is a subspace of the first.

If A and B belong to $\mathfrak{N}$ and are similar then their tangent spaces defined above are related by the appropriate similarity. Further note that $C \in \tan(J, A)$ if and only if $(d/dt)E_r(A + tC)\,|_{t=0} = 0$ for all $r = 1, 2, \ldots, n$. If $A \in \mathfrak{N}$ is of nilindex n, then, by taking A into upper Jordan canonical form, one sees that the equations for $X \in \tan(J, A)$ are, up to a similarity,

$$0 = \sum_{i=0}^{n-j} X_{j+i,i+1}, \qquad j = 1, 2, \ldots, n.$$

Therefore $\dim \tan(J, A) = n^2 - n$. Since J is generated by n polynomials, if N is of nilindex n we have [7, p. 28, 37]

$$n^2 - n \leq \dim \mathfrak{N} \leq \dim \tan(\operatorname{rad} J, N) \leq \dim \tan(J, N) = n^2 - n.$$

So if N is of nilindex n then $\tan(\operatorname{rad} J, N) = \tan(J, N)$.

LEMMA 2. *If $A, B \in \mathfrak{N}$ are both of nilindex n then $AB = BA$ if and only if* $\tan(\operatorname{rad} J, A) = \tan(\operatorname{rad} J, B)$.

Proof. A is of nilindex n so its minimal and characteristic polynomials are equal. Therefore, if $AB = BA$, then B is a polynomial in A. By the above remarks, we may assume that

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

so

$$B = \begin{bmatrix} 0 & a_1 & a_2 & \cdots & a_{n-1} \\ 0 & 0 & a_1 & \cdots & a_{n-2} \\ 0 & 0 & 0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a_1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

where $a_i \in K$. Since B is of nilindex n, $a_1 \neq 0$. A direct computation shows that

$$\frac{d}{dt} E_n(B + tX)\Big|_{t=0} = a_1^{n-1} X_{n1}.$$

Hence the equation for B arising from E_n is $X_{n1} = 0$, which is the same as for A. One has that

$$\frac{d}{dt} E_r(B + tX)\Big|_{t=0} = a_1^{r-1} \sum_{i=0}^{n-r} X_{r+i\,i+1} + \sum_{j=1}^{r-1} A_j \sum_{i=0}^{n-j} X_{j+i\,i+1}$$

for suitable constants A_j depending on $a_1,\ldots,a_{n-1}$. By induction, the equation for B arising from E_r is

$$a_1^{r-1}\sum_{i=0}^{n-r} X_{r+i\,i+1} = 0,$$

and since $a_1 \neq 0$ this is the same as for A. Since $\tan(J, A) = \tan(\operatorname{rad} J, A)$ the results follows.

On the other hand, suppose $\tan(\operatorname{rad} J, A) = \tan(\operatorname{rad} J, B)$. We may assume A is as above. Let E_{ij} be the matrix with 1 in the (i, j) position and zeros elsewhere. Then

$$E_{ji} \in \tan(\operatorname{rad} J, A), \qquad i > j,$$

and

$$E_{ji} - E_{j+1\,i+1} \in \tan(\operatorname{rad} J, A), \qquad i \leq j.$$

Writing $B = (b_{ij})$, we have

$$\frac{d}{dt}E_2(B + tE_{ji})\Big|_{t=0} = b_{ij}, \qquad \text{if } i > j,$$

$$\frac{d}{dt}E_2(B + t(E_j i - E_{j+1\,i+1}))\Big|_{t=0} = \pm(b_{ij} - b_{i+1\,j+1}) \qquad \text{if } i \leq j.$$

Therefore $b_{ij} = 0$ if $i > j$ and $b_{ij} = b_{i+1,j+1}$ if $i \leq j$, and B is a polynomial in A.

LEMMA 3. *If L: $\mathrm{sl}_n \to \mathrm{sl}_n$ is a nonsingular linear transformation with the property that $L(\mathfrak{N}) = \mathfrak{N}$, and $A \in \mathfrak{N}$, then* $L(\tan(\operatorname{rad} J, A)) = \tan(\operatorname{rad} J, L(A))$.

Proof. The map $\tilde{L}$: $K[X] \to K[X]$ defined by $\tilde{L}(f)(A) = f(L(A))$ is a K-algebra homomorphism. Since L is nonsingular and $L(\mathfrak{N}) = \mathfrak{N}$ and $\operatorname{rad} J = \{f \in K[X]\colon f(N) = 0, \text{ for all } N \in \mathfrak{N}\}$, we have $\tilde{L}(\operatorname{rad} J) = \operatorname{rad} J$. Thus

$$\begin{aligned}
\tan(\operatorname{rad} J, L(A)) &= \left\{B \in M_n\colon \frac{df}{dt}(L(A) + tB)\Big|_{t=0} \quad \text{for all } f \in \operatorname{rad} J\right\} \\
&= \left\{L(C) \in M_n\colon \frac{df}{dt}(L(A) + tL(C))\Big|_{t=0} \quad \text{for all } f \in \operatorname{rad} J\right\} \\
&= \left\{L(C) \in M_n\colon \frac{d\tilde{L}(f)}{dt}(A + tC)\Big|_{t=0} = 0 \quad \text{for all } f \in \operatorname{rad} J\right\} \\
&= L(\tan(\operatorname{rad} J, A)).
\end{aligned}$$

Proof of theorem. First we observe that $L(\mathfrak{N}) = \mathfrak{N}$. This follows from Lemma 1 of Dixon [5] and the fact that L is nonsingular.

We now show that L preserves nilindex n. If $A \in \mathfrak{N}$ and rank $A \leq n - 2$, then A kills two linearly independent vectors v, w. Let $\mathfrak{M}_1$, $\mathfrak{M}_2$ be maximal nilpotent algebras containing A and killing v, w respectively. Every maximal nilpotent algebra kills exactly one line, so $\mathfrak{M}_1 \neq \mathfrak{M}_2$. By Lemma 1, L maps maximal nilpotent algebras to maximal nilpotent algebras and again by lemma 1, L preserves the matrices of nilindex n.

Now we show that if $A \in \mathfrak{N}$ has rank one, then so does $L(A)$. Let U be the unit auxiliary matrix $E_{12} + \cdots + E_{n-1,n}$.

First note that the only members of $\mathfrak{U}$ which commute with U and E_{12} are multiples of E_{1n}. Thus the centre of any maximal nilpotent algebra is one-dimensional and is generated by a rank one matrix.

Let $A \in \mathfrak{N}$ have rank one. Then for some nonsingular S, $S^{-1}AS = E_{1n}$. Let $\mathfrak{M} = S\mathfrak{U}S^{-1}$. Then A generates the centre of $\mathfrak{M}$. Let $V \in \mathfrak{M}$ have nilindex n. Then V and $A + V$ have nilindex n and commute. It follows from Lemmas 2 and 3 that $L(A + V)$ commutes with $L(V)$. Hence $L(A)$ commutes with $L(V)$. Since the nilindex n matrices in $\mathfrak{M}$ generate $\mathfrak{M}$, $L(A)$ is in the centre of the maximal nilpotent algebra $L(\mathfrak{M})$. Hence $L(A)$ has rank one.

We next define two bijections on the lines (through the origin) of K^n and use the fundamental theorem of projective geometry. For each line $\langle v \rangle \in K^n$, define two $n - 1$ dimensional subspaces of $\mathfrak{N}$ by

$$M(v) = \{X \in \mathfrak{N} \mid \operatorname{Im} X = \langle v \rangle\},$$

$$M^t(v) = \{X^t \mid X \in M(v)\}.$$

We will show that $L(M(v)) = M(w)$ or $M^t(w)$ and $L(M^t(v)) = M(w')$ or $M^t(w')$ for some $w, w' \in K^n$. The bijections will be $\varphi(v) = w$ and $\theta(v) = w'$.

We note a few facts about $M(v)$. Any nonzero member of $M(v)$ has rank one. If $v, w \in K^n$, and are nonzero, then $M(v)$ and $M(w)$ are conjugate, and if $w = Av$, A nonsingular, then $M(w) = AM(v)A^{-1}$. In tensor notation, $M(v) = v \otimes v^{\perp}$ and $M^t(v) = v^{\perp} \otimes v$. (Here, $\perp$ means orthogonal complement with respect to the dot product.) It is easily verified that $M(u) \cap M(v) = M(u) = M(v)$ if u and v are linearly dependent and 0 otherwise, and that $M(u) \cap M^t(v) = \langle u \otimes v \rangle$ if $u \cdot v = 0$ and is 0 otherwise. Finally, observe that any $n - 1$ dimensional subspace of $\mathfrak{N}$ with all of its nonzero matrices having rank one must be an $M(v)$ or an $M^t(v)$. It follows that for $v \in K^n$, there is a $w \in K^n$ such that $L(M(v)) = M(w)$ or $M^t(w)$.

Suppose we have $v, w \in K^n$ with $L(M(v)) = M(v')$ and $L(M(w)) = M^t(w')$. Since $n \geq 3$, pick u orthogonal to v and w. Then $M(v) \cap M^t(u)$

and $M(w) \cap M^t(u)$ are one dimensional. If $L(M^t(u)) = M(u')$ then $M(u') \cap M(v') = L(M^t(u) \cap M(v))$ has dimension 1; which is impossible, as $M(u') \cap M(v')$ has dimension 0 or $n - 1 \geq 2$. If $L(M^t(u)) = M^t(u')$, we reach a similar contradiction. A similar argument holds when we examine the images of $M^t(v)$ and $M^t(w)$. Thus, by replacing L with the map $X \to L(X)^t$ if necessary, we may assume that for any nonzero $v \in K^n$, $L(M(v)) = M(w)$ and $L(M^t(v)) = M^t(u)$ for appropriate $u, w \in K^n$.

Thus we define two maps φ, θ induced by L on the lines of K^n. We have $L(M(v)) = M(\varphi(v))$ and $L(M^t(v)) = M^t(\theta(v))$ for $v \in K^n$.

Since $L(\mathcal{N}) = \mathcal{N}$, L^{-1} also preserves nilpotence and hence φ and θ are bijections on the lines of K^n.

Now we show that φ and θ preserve coplanarity of lines in K^n and thus satisfy the hypothesis of the fundamental theorem of projective geometry. Let $\langle u_1 \rangle$, $\langle u_2 \rangle$, $\langle u_3 \rangle$ be three distinct coplanar lines in K^n. Then

$$\begin{aligned} 2n - 1 &= \dim(M(u_1) + M(u_2) + M(u_3)) \\ &= \dim L(M(u_1) + M(u_2) + M(u_3)) \\ &= \dim(M(\varphi(u_1)) + M(\varphi(u_2)) + M(\varphi(u_3))). \end{aligned}$$

If $\varphi(u_1)$, $\varphi(u_2)$, $\varphi(u_3)$ are linearly independent then

$$\dim(M(\varphi(u_1)) + M(\varphi(u_2)) + M(\varphi(u_3))) = 3n - 3$$

and this is impossible since $n \geq 3$. Thus $\varphi(u_1)$, $\varphi(u_2)$, $\varphi(u_3)$ are coplanar and φ satisfies the hypothesis of the fundamental theorem of projective geometry. So does θ. Thus there exist semilinear maps S and T on K^n such that $\varphi(u) = \langle Su \rangle$ and $\theta(u) = \langle Tu \rangle$, for all nonzero u in K^n.

There are linear maps P and Q on K^n and automorphisms σ and τ on K such that $Sv = P(\sigma v)$ and $Tv = Q(\tau v)$. (The automorphisms act componentwise.) Then

$$L(M(v)) = M(P\sigma v) = PM(\sigma v)P^{-1}$$

and

$$L(M^t(v)) = M^t(Q\tau v) = Q^{t^{-1}}M^t(\tau v)Q^t.$$

Suppose $u \cdot v = 0$. Then $\dim(M(u) \cap M^t(v)) = 1$ and so

$$\dim(M(P\sigma u) \cap M^t(Q\tau v)) = 1$$

and thus $(P\sigma u) \cdot (Q\tau v) = 0$, i.e.,

$$u \cdot \sigma^{-1}(P^tQ\tau v) = 0.$$

Let R be the semilinear map defined by

$$Rv = \sigma^{-1}(P^tQ\tau v).$$

Then $u \cdot v = 0$ implies $u \cdot Rv = 0$. Thus $R = dI$ is a scalar map, $\sigma = \tau$ and $P^tQ = dI$.

Replace the map L with the map $X \to P^{-1}L(X)P$. Then $L(M(v)) = M(\sigma v)$ and $L(M^t(v)) = M^t(\sigma v)$, for all nonzero v in K^n. Thus if $u \cdot v = 0$ then $L(u \otimes v) = c(u \otimes v)\sigma(u \otimes v)$, where c is a scalar valued function. If $v \in \langle u_1, u_2 \rangle^\perp$, then by comparing $L((u_1 + u_2) \otimes v)$ with $L(u_1 \otimes v) + L(u_2 \otimes v)$ we get $c(u_1 \otimes v) = c(u_2 \otimes v)$. Similarly if $u \in \langle v_1, v_2 \rangle$, then $c(u \otimes v_1) = c(u \otimes v_2)$.

Now we show that c is a constant function. Suppose that $u_1 \cdot v_1 = 0$ and $u_2 \cdot v_2 = 0$. Pick $v_3 \in \langle u_1, u_2 \rangle^\perp$. Then $c(u_1 \otimes v_1) = c(u_1 \otimes v_3) = c(u_2 \otimes v_3) = c(u_2 \otimes v_2)$. Thus c is a constant function say $c(u \otimes v) = k$. Then $L(u \otimes v) = k\sigma(u \otimes v)$, for all u, v with $u \cdot v = 0$.

Since L is linear, σ is the identity automorphism. The rank one nilpotent matrices span sl_n and so the theorem is proved.

REMARK. When $n = 2$, the same result is obtained by a simple computation.

REFERENCES

1. E. Artin, *Geometric Algebra*, Interscience, New York, 1957.

2. E. P. Botta and S. Pierce, *The preservers of any orthogonal group*, Pacific J. Math., **70** (1977), 37–50.

3. C. Chevalley, *Théorie des Groupes de Lie*, Hermann, Paris, 1968.

4. J. Dieudonné, *Sur une généralisation du groupe orthogonal à quatre variables*, Arch. Math., **1** (1948), 282–287.

5. J. Dixon, *Rigid embeddings of simple groups in the general linear group*, Canad. J. Math., **29** (1977), 384–391.

6. M. Gerstenhaber, *On nilalgebras and linear varieties of nilpotent matrices I*, Amer. J. Math., **80** (1958), 614–622.

7. J. E. Humphreys, *Linear Algebraic Groups*, Springer-Verlag, New York, 1975.

8. H. G. Jacob, *Coherence invariant mappings on Kronecker products*, Amer. J. Math., **77** (1955), 177–189.

9. M. Marcus, *Linear transformations on matrices*, J. Res., Nat. Bur. of Stand., v. 75B, Nos. 3 and 4 (1971), 107–113.

10. M. Marcus and B. Moyls, *Transformations on tensor product spaces*, Pacific J. Math., **9** (1959), 1215–1221.

Received March 16, 1979 and in revised form December 28, 1980. The work of the second author was partially supported by NCR Grant A-7862.

UNIVERSITY COLLEGE
TORONTO M5S 1A1 CANADA

UNIVERSITY COLLEGE
TORONTO M5S 1A1 CANADA

AND

CALIFORNIA STATE UNIVERSITY
NORTHRIDGE, CA 91330