A NOTE ON THE REPLICAS OF NILPOTENT MATRICES

HSIO-FU TUAN

In a recent paper,[1] Chevalley proved the following theorem:

(A) *If Z is a nilpotent matrix over a field K of characteristic* 0, *the only replicas Z' of Z are the matrices $Z'=tZ$, $t\in K$.*[2]

For the proof of (A), he made use of a particular case of a theorem due to Ado and gave a proof for the results which he needed. In the present note, we shall give a direct simple proof of (A) and we shall in fact deduce it as an immediate consequence of the stronger theorem:

(B) *If Z and Z' are two nilpotent matrices over a field K of characteristic* 0, *and if $q(x)$ and $r(x)$ are two polynomials with coefficients in K and without constant terms such that $Z'=q(Z)$ and $Z'_{0,2}=r(Z_{0,2})$, then $Z'=tZ$, $t\in K$.*

We shall later establish corresponding results for fields K of prime characteristics, to be stated as theorems (C) and (D).

That (A) is implied by (B) follows immediately from the fact that if Z' is a replica of Z, then $Z'_{r,s}=p_{r,s}(Z_{r,s})$, where $p_{r,s}(x)$ are polynomials in K without constant terms.[3]

For the proof of (B), let n be the degree of Z and Z' and let m be the least nonnegative integer such that $Z^{m+1}=0$. Clearly $0\leqq m\leqq n-1$. The case $Z=0$ is trivial; we can therefore assume $1\leqq m\leqq n-1$. Let also l be the least nonnegative integer such that $(Z_{0,2})^{l+1}=0$. Clearly $Z_{0,2}$ is nilpotent and $1\leqq l\leqq n^2-1$. We shall see that $m\leqq l\leqq 2m$.

The matrix Z can be transformed by an (n, n) matrix T with coefficients in the algebraic closure $\overline{K}$ of K into the following form:

$$Z_1 = T^{-1}ZT = \begin{pmatrix} 0 & & & & \\ z_1 & 0 & & & \\ & \ddots & \ddots & & \\ & & \ddots & 0 & \\ & & & z_{n-1} & 0 \end{pmatrix}, \tag{1}$$

where $z_1, \cdots, z_{n-1}$ are zeros and ones and not all zeros. Then for

Received by the editors July 17, 1944.

[1] Claude Chevalley, *On a kind of new relationship between matrices*, Amer. J. Math. vol. 65 (1943) pp. 521–531.

[2] Theorem 6, p. 530, loc. cit.

[3] Lemma 4, p. 529, loc. cit.

any integer i, $1 \leqq i \leqq m$, we have

$$(1)_i \qquad Z_1^i = T^{-1} Z^i T = \begin{pmatrix} 0 & & & & & \\ \vdots & \ddots & & & & \\ 0 & & \ddots & & & \\ z_{i,1} & \ddots & & \ddots & & \\ & \ddots & \ddots & & \ddots & \\ & & z_{i,j} & \ddots & & \ddots \\ & & & \ddots & \ddots & \\ & & & & z_{i,n-i} & \underbrace{0 \cdots 0}_{i} \end{pmatrix},$$

$$z_{i,j} = z_j \cdots z_{i+j-1}, \qquad j = 1, \cdots, n-i.$$

Hence if we write

$$(2) \qquad q(x) = q_1 x + \cdots + q_m x^m, \; q_i \in K, \quad i = 1, \cdots, m,$$

we have

$$(3) \qquad Z_1' = T^{-1} Z' T = T^{-1} q(Z) T = q(T^{-1} Z T) = q(Z_1).$$

Denoting the (n, n) identity matrix by E, we have then[4]

$$(4)_1 \qquad (Z_1)_{0,2} = Z_1 \oplus Z_1 = Z_1 \otimes E + E \otimes Z_1$$

$$= \begin{pmatrix} 0 & & & & \\ z_1 E & 0 & & & \\ & \ddots & \ddots & & \\ & & \ddots & 0 & \\ & & & z_{n-1} E & 0 \end{pmatrix} + \begin{pmatrix} Z_1 & & & & \\ & Z_1 & & & \\ & & \ddots & & \\ & & & Z_1 & \\ & & & & Z_1 \end{pmatrix}$$

$$= \begin{pmatrix} Z_1 & & & & \\ z_1 E & Z_1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & Z_1 & \\ & & & z_{n-1} E & Z_1 \end{pmatrix},$$

hence for any positive integer i,[5]

[4] We define $A \otimes B = (a_{ij}) \otimes B = (a_{ij} B)$. Observe that $(A \otimes B)(C \otimes D) = AB \otimes CD$. The heavy cross and heavy plus used in Chevalley's paper are here replaced by $\otimes$ and $\oplus$.

[5] In the following, * denotes the terms in which we are not interested.

$$((Z_1)_{0,2})^{i} = (Z_1 \otimes E + E \otimes Z_1)^{i}$$

$$= Z_1 \otimes E + \sum_{j=1}^{i-1} C_{i,j} Z_1^{i-j} \otimes Z_1^{j} + E \otimes Z_1^{i}$$

$$(4)_i \qquad = \begin{pmatrix} Z_1^{i} & & & & \\ & Z_1^{i} & & & \\ & & \ddots & & \\ & * & & Z_1^{i} & \\ & & & & Z_1^{i} \end{pmatrix}$$

($C_{i,j}$ being the binomial coefficients), and therefore $m \leqq l \leqq 2m$. We may write

$$(5) \qquad r(x) = r_1 x + \cdots + r_l x^l, \quad r_i \in K, \qquad i = 1, \cdots, l,$$

and then

$$(6) \qquad \begin{aligned} (Z_1')_{0,2} &= Z_1' \oplus Z_1' = Z_1' \otimes E + E \otimes Z_1' \\ &= T^{-1}Z'T \otimes T^{-1}ET + T^{-1}ET \otimes T^{-1}Z'T \\ &= (T^{-1} \otimes T^{-1})(Z' \otimes E + E \otimes Z')(T \otimes T) \\ &= (T \otimes T)^{-1} Z_{0,2}' (T \otimes T) = (T \otimes T)^{-1} r(Z_{0,2})(T \otimes T) \\ &= r((T \otimes T)^{-1} Z_{0,2} (T \otimes T)) = r((Z_1)_{0,2}). \end{aligned}$$

Consequently the same relations originally assumed for Z and Z' now hold for Z_1 and Z_1'. For simplicity in notations, we shall now just consider Z and Z' for Z_1 and Z_1' in the related formulas $(1)_i$, (3), $(4)_i$, (6).

Now, on the one hand,

$$(7) \qquad \begin{aligned} Z_{0,2}' &= Z' \oplus Z' = Z' \otimes E + E \otimes Z' = q(Z) \otimes E + E \otimes q(Z) \\ &= E \otimes q(Z) + (q_1 Z + \cdots + q_m Z^m) \otimes E \\ &= E \otimes q(Z) + Z \otimes q_1 E + \cdots + Z^m \otimes q_m E \\ &= \begin{pmatrix} q(Z) & & & & & \\ q_1 z_1 E & q(Z) & & & & \\ & q_2 z_2 E & q(Z) & & & \\ & * & \ddots & \ddots & & \\ & & & & q(Z) & \\ & & & & q_1 z_{n-1} E & q(Z) \end{pmatrix}, \end{aligned}$$

while, on the other hand,

$$
\begin{aligned}
Z'_{0,2} &= r(Z_{0,2}) = r_1 Z_{0,2} + r_2 (Z_{0,2})^2 + \cdots + r_l (Z_{0,2})^l \\
&= r_1(Z \oplus Z) + r_2(Z \oplus Z)^2 + \cdots + r_l(Z \oplus Z)^l \\
&= r_1(Z \otimes E + E \otimes Z) + r_2(Z^2 \otimes E + 2Z \otimes Z + E \otimes Z^2) \\
&\quad + \cdots + r_l\left(Z^l \otimes E + \sum_{i=1}^{l-1} Z^{l-i} \otimes Z^i + E \otimes Z^l\right) \\
&= E \otimes (r_1 Z + r_2 Z^2 + \cdots + r_l Z^l) \\
&\quad + Z \otimes (r_1 E + 2r_2 Z + \cdots + l r_l Z^{l-1}) \\
&\quad + \cdots + Z^l \otimes r_l E \\
&= E \otimes r(Z) + Z \otimes r'(Z) + Z^2 \otimes (1/2) r''(Z) \\
&\quad + \cdots + Z^l \otimes (1/l!) r^{(l)}(Z) \\
&= \begin{pmatrix}
r(Z) & & & & & \\
z_1 r'(Z) & r(Z) & & & & \\
 & z_2 r'(Z) & r(Z) & & & \\
 & & \ddots & \ddots & & \\
 & * & & \ddots & r(Z) & \\
 & & & & z_{n-1} r'(Z) & r(Z)
\end{pmatrix},
\end{aligned} \tag{8}
$$

where $r'(x), r''(x), \cdots, r^{(l)}(x)$ are the successive derivatives of $r(x)$.

In (7) and (8), comparing the terms (which are (n, n) matrices) on the main diagonal and on the first parallel just below, we obtain

$$q(Z) = r(Z), \tag{9}$$

$$q_1 z_i E = z_i r'(Z); \qquad i = 1, \cdots, n-1. \tag{10}$$

(9) gives

$$q_1 Z + \cdots + q_m Z^m = r_1 Z + \cdots + r_l Z^l, \qquad l \geqq m,$$

and hence

$$q_j = r_j, \qquad j = 1, \cdots, m. \tag{11}$$

(10) gives

$$q_1 z_i E = z_i (r_1 E + 2r_2 Z + \cdots + l r_l Z^{l-1}), \qquad i = 1, \cdots, n-1,$$

by (11),

$$z_i (2q_2 E + \cdots + m q_m Z^{m-1} + \cdots) = 0, \qquad i = 1, \cdots, n-1,$$

as not all z_i (for $i=1, \cdots, n-1$) are zero, hence

$$2q_2 E + \cdots + m q_m Z^{m-1} + \cdots = 0,$$

and consequently

$$2q_2 = \cdots = mq_m = 0. \tag{12}$$

Now K is of characteristic 0; we can therefore conclude

$$q_2 = \cdots = q_m = 0. \tag{13}$$

It then follows that

$$Z' = q_1 Z = tZ \tag{14}$$

with $q_1 = t \in K$, as is to be proved.

Let us now suppose that K is of prime characteristic, say p. If $p > m$, which is certainly the case if $p \geqq n$, then we can still infer (13) from (12) and hence still have (14) as before. For the cases $p \leqq m$, from (12) we can only infer that

$$q_i = 0, \qquad 2 \leqq i \leqq m, \qquad p \nmid i, \tag{15}$$

and hence we can only conclude that Z' is necessarily of the form

$$Z' = q_1 Z + q_p Z^p + q_{2p} Z^{2p} + \cdots + q_{m'p} Z^{m'p}, \tag{16}$$

where $m' = [m/p]$, the largest integer not greater than m/p. In fact, the conclusions (13) and (14) are then no longer true in general.

We shall first prove that $Z' = Z^\alpha$ with $\alpha = p^a$ (a being any non-negative integer) is a replica of Z. This follows from the facts:

$$\begin{aligned} Z'_{r,s} &= (Z^\alpha)_{r,s} = \underbrace{(Z^\alpha) \oplus \cdots \oplus (Z^\alpha)}_{r} \oplus \underbrace{(Z^\alpha) \oplus \cdots \oplus (Z^\alpha)}_{s} \\ &= (-{}^tZ^\alpha) \oplus \cdots \oplus (-{}^tZ^\alpha) \oplus Z^\alpha \oplus \cdots \oplus Z^\alpha \\ &= -({}^tZ^\alpha) \otimes \cdots \otimes E \otimes E \otimes \cdots \otimes E - \cdots \\ &\quad - E \otimes \cdots \otimes ({}^tZ^\alpha) \otimes E \otimes \cdots \otimes E \\ &\quad + E \otimes \cdots \otimes E \otimes Z^\alpha \otimes \cdots \otimes E + \cdots \\ &\quad + E \otimes \cdots \otimes E \otimes E \otimes \cdots \otimes Z^\alpha \\ &= (-{}^tZ \otimes \cdots \otimes E \otimes E \otimes \cdots \otimes E - \cdots \\ &\quad - E \otimes \cdots \otimes {}^tZ \otimes E \otimes \cdots \otimes E \\ &\quad + E \otimes \cdots \otimes E \otimes Z \otimes \cdots \otimes E + \cdots \\ &\quad + E \otimes \cdots \otimes E \otimes E \otimes \cdots \otimes Z)^\alpha \\ &= (Z_{r,s})^\alpha, \end{aligned} \tag{17}$$

because $(-1)^\alpha = \pm 1 \equiv -1 \pmod{p}$ for $p = 2$ and $(-1)^\alpha = -1$ for $p \neq 2$.

We shall now prove that conversely if $Z'=q(Z)=\sum_{i=1}^{m}q_iZ^i$ with $q_i\in K$ is a replica of Z, then we have

$$(18)\quad Z' = q(Z) = \sum_{j=1}^{m_0} q_{i(j)}Z^{i(j)}, \qquad i(j) = p^j, \quad i(m_0) \leqq m < i(m_0+1).$$

We shall show more strongly that only then $Z'_{1,1}=s(Z_{1,1})$, where $s(x)$ is a polynomial with coefficients in K and without the constant term.

We can assume that Z is of the form (1); because as before,

$$(Z'_1)_{1,1} = (T^{-1}Z'T)_{1,1} = ({}^tT^{-1}\otimes T)^{-1}Z'_{1,1}({}^tT^{-1}\otimes T),$$

hence $Z'_{1,1}=r(Z_{1,1})$ implies also

$$(19)\quad \begin{aligned}(Z_1')_{1,1} &= ({}^tT^{-1}\otimes T)^{-1}s(Z_{1,1})({}^tT^{-1}\otimes T)\\ &= s(({}^tT^{-1}\otimes T)^{-1}Z_{1,1}({}^tT^{-1}\otimes T)) = s((Z_1)_{1,1}).\end{aligned}$$

Then, for any positive integer i, $1\leqq i\leqq m$, we have

$$(20)_i\quad \begin{aligned}(Z')^i &= (-{}^tZ)^i = (-1)^i\,{}^tZ_i\\ &= \begin{pmatrix} 0 & \cdots & 0 & z_{i,1} & & & \\ & \ddots & & \ddots & \ddots & & \\ & & \ddots & & z_{i,j} & & \\ & & & \ddots & & \ddots & \\ & & & & \ddots & & z_{i,n-i} \\ & & & & & \ddots & 0 \\ & & & & & & \vdots \\ & & & & & & 0 \end{pmatrix},\end{aligned}$$

$$z_{i,j} = (-1)^i z_j\cdots z_{i+j-1}, \qquad j = 1,\cdots,n-i,$$

and further

$$(21)_i\quad \begin{aligned}(Z_{1,1})^i &= (Z'\oplus Z)^i = (Z'\otimes E+E\otimes Z)^i = (-{}^tZ\otimes E+E\otimes Z)^i\\ &= (-1)^i\,{}^tZ^i\otimes E+\sum_{j=1}^{i-1}(-1)^{i-j}C_{i,j}Z^{i-j}\otimes Z^j+E\otimes Z^i\\ &= \begin{pmatrix} Z & -z_1E & & & \\ & Z & \ddots & & \\ & & \ddots & \ddots & \\ & & & Z & -z_{n-1}E \\ & & & & Z\end{pmatrix}^i = \begin{pmatrix} Z^i & & & & \\ & Z^i & & * & \\ & & \ddots & & \\ & & & Z^i & \\ & & & & Z^i\end{pmatrix},\end{aligned}$$

hence it follows as before that the least nonnegative integer k such that $(Z_{1,1})^{k+1}=0$ satisfies $m\leqq k\leqq 2m$. We can therefore write

$$
\begin{aligned}
s(x) &= s_1x + \cdots + s_kx^k \qquad (s_i \in K) \\
&= (s_1x + \cdots + s_mx^m) + (s_{m+1}x^{m+1} + \cdots + s_kx^k) \qquad (22)\\
&= s'(x) + s''(x).
\end{aligned}
$$

We have

$$
\begin{aligned}
Z'_{1,1} &= Z' \oplus Z = (-{}^tZ') \oplus Z = -{}^tZ' \otimes E + E \otimes Z' \\
&= -q({}^tZ) \otimes E + E \otimes q(Z) \qquad (23)\\
&= E \otimes q(Z) + \sum_{j=1}^{m} {}^tZ^j \otimes (-q_jE),
\end{aligned}
$$

while, on the other hand,

$$
\begin{aligned}
Z'_{1,1} &= \sum_{i=1}^{k} s_i(Z_{1,1})^i = \sum_{i=1}^{k} s_i(-{}^tZ \otimes E + E \otimes Z)^i \\
&= \sum_{i=1}^{k} s_i \sum_{j=\max(0,i-m)}^{\min(i,m)} C_{i,j}(-{}^tZ)^j \otimes Z^{i-j} \\
&= \sum_{j=0}^{m} {}^tZ^j \otimes (-1)^j \left(\sum_{i=\max(1,j)}^{\min(k,m+j)} s_iC_{i,j}Z^{i-j} \right) \qquad (24)\\
&= E \otimes \sum_{i=1}^{m} s_iZ_i + \sum_{j=1}^{m} {}^tZ^j \otimes (-1)^j \left(\sum_{i=j}^{\min(k,m+j)} s_iC_{i,j}Z^{i-j} \right) \\
&= E \otimes s'(Z) + \sum_{j=1}^{m} {}^tZ^j \otimes s_j(Z),
\end{aligned}
$$

where $s_j(x)$ are polynomials with coefficients in K for $j=1, \cdots, m$ (observe $\max(0, i-m)\leqq \min(i, k-m)\leqq \min(i, m)$). Writing the matrix $Z'_{1,1}$ as given by (23) and (24) in the form of two compound (n, n) matrices whose elements are again (n, n) matrices and comparing the terms on their main diagonals and on their m first parallels above the main diagonals, we can then conclude that first

$$
q(Z) = \sum_{i=1}^{m} q_iZ^i = \sum_{i=1}^{m} s_iZ^i = s(Z), \qquad (25)
$$

$$
q_i = s_i, \qquad i = 1, \cdots, m; \qquad (26)
$$

and that also

$$
z_{j,h}(-q_jE) = z_{j,h}s_j(Z), \qquad j = 1, \cdots, m,\ h = 1, \cdots, n-j,
$$

as $Z'^{j}\neq 0$ for $j=1, \cdots, m$, so for each j we have at least one h, say $h(j)$ (=one of $1, \cdots, n-j$), such that $z_{j,h(j)}\neq 0$, then $z_{j,h(j)}=\pm 1$, and hence

$$- q_jE = s_j(Z) = (-1)^j \sum_{i=j}^{\min(k,m+j)} s_iC_{i,j}Z^{i-j}$$

$$= (-1)^j s_jE + (-1)^j \sum_{i=j+1}^{\min(k,m+j)} s_iC_{i,j}Z^{i-j},^{6}$$

$$0 = (1 + (-1)^j)s_jE + (-1)^j \sum_{i=j+1}^{\min(k,m+j)} s_iC_{i,j}Z^{i-j}, \tag{27}$$

$$(1 + (-1)^j)s_j = 0, \qquad s_iC_{i,j} = 0, \qquad j = 1, \cdots, m; \quad i = j+1, \cdots, \min(k, m+j)^{6} \tag{28}$$

(observe $\min(k, m+j)\geqq m$), in particular

$$s_iC_{i,j} = 0, \qquad i = 2, \cdots, m; j = 1, \cdots i-1. \tag{29}$$

Now it is easily seen that if $i=p^ji'$ with $p\nmid i'$, then $p\nmid C_{i,i/i'}$,[7] hence it follows immediately that if $i'\neq 1$, namely, $i\neq p^j$ for $j=1, \cdots, m_0$ with $p^{m_0}\leqq m<p^{m_0+1}$, then $q_i=s_i=0$ for $i=2, \cdots, m$, as was to be proved.

Summarizing our results for fields K of characteristic $p\neq 0$, we have the following theorems:

(C) *If Z and Z' are two nilpotent matrices over a field K of characteristic $p\neq 0$, and if $q(x)$, $r(x)$ and $s(x)$ are three polynomials with coefficients in K and without constant terms such that $Z'=q(Z)$, $Z'_{0,2}=r(Z_{0,2})$ and $Z'_{1,1}=s(Z_{1,1})$, then we have*

$$Z' = q(Z) = \sum_{j=0}^{m_0} t_jZ^{i(j)}, \qquad i(j) = p^j, \qquad t_j \in K. \tag{30}$$

(D) *If Z is a nilpotent matrix over a field K of characteristic $p\neq 0$, the only replicas Z' of Z are the matrices* (30).

PRINCETON UNIVERSITY

---

[6] If $k=m$, then for $j=m$, there is no i, $j+1\leqq i\leqq \min(k, m+j)$.

[7] Consider $p^j=p^j \cdots 1$ and $i\cdots(i-p^j+1)=(p^j+p^j(i'-1))\cdots(1+p^j(i'-1))$. As $p^v\|u$ (namely $p^v|u$ and $p^{v+1}\nmid u$) for $u=p^j, \cdots, 1$ implies $p^v\|(u+p^j(i'-1))$ because $v\leqq j$, so $p\nmid C_{i,i/i'}$.