GROUP INVERSES AND DRAZIN INVERSES OF BIDIAGONAL AND TRIANGULAR TOEPLITZ MATRICES

R. E. HARTWIG and J. SHOAF

Communicated by M. F. Newman

Abstract

Necessary and sufficient conditions are given for the existence of the group and Drazin inverses of bidiagonal and triangular Toeplitz matrices over an arbitrary ring.

Introduction

In the study of matrix equations and generalised inverses (Hartwig (1975)), the question arises when the bidiagonal hypercompanion block matrix

$$M = \begin{bmatrix} L & & & & 0 \\ N & L & & & \\ & N & \cdot & & \\ & & \cdot & \cdot & \\ 0 & & & N & L \end{bmatrix}$$

possesses a group inverse (Ben-Israel and Greville, (1974)); that is when does there exist a solution $M^{\#}$ to $MXM = M$, $XMX = X$, $MX = XM$. The purpose of this paper is to answer this question in the most general setting and to give an expression for $M^{\#}$ in terms of L and N, when it exists. Since the concept of a Drazin inverse (Ben-Israel and Greville (1974), Drazin (1958)), of a ring element is closely related to that of a group inverse, we shall at the same time investigate the existence of this type of generalised inverse for M.

The following results from ring theory (Hartwig (1976)), Proposition 7, will be assumed. *If a is an element of a ring R, then there exists a solution $a^{\#}$ to*

10

(1.1)   (1) $axa = a$,   (2) $xax = x$,   (3) $ax = xa$

if and only if there are solutions to $a^2u = a$, $a = va^2$. In either case $a^* = vau$ is unique and

$$a^0 = \{x \mid ax = 0\} = (a^2)^0 \qquad {}^0a = \{x \mid xa = 0\} = {}^0(a^2).$$

In the language of generalised inverses the solutions to (1.1)-1 and (1.1)-2 are called 1-inverses and 2-inverses of a respectively (Ben-Israel and Greville (1974)). The Drazin inverse a^d (Drazin (1958)) of an element $a \in R$ is the unique solution, if any, to

(1.2) $a^kxa = a^k$, $xax = x$ $ax = xa$,

and exists if and only if there are non-negative integers p and q for which there exist solutions to $a^{p+1}u = a^p$, and $va^{q+1} = a^q$. The smallest values of p, q and k for which there exist such solutions, if any, are called the left index $l(a)$, the right index $r(a)$ and the index $i(a)$ respectively. It was shown in Drazin (1958) that if $l(a)$ and $r(a)$ are finite then $l(a) = r(a) = i(a)$. It should further be noted that if $i(a) = 0$, or 1, then $a^d = a^*$, and that a^d exists if and only if $(a^k)^*$ exists for some positive integer k (so that $k \geq i(a)$) and is then given by

(1.3) $a^d = a^{k-1}(a^k)^* = (a^k)^*a^{k-1}$.

Throughout this paper we denote the ring of $n \times n$ matrices over R by $R_{n \times n}$, and we shall assume for notational convenience that our rings have unity even though most of our results do not depend upon this fact.

1. Preliminary Results

We begin with two preliminary results.

LEMMA 1.   *Let R be a ring with unity 1 and let $u, x \in R$ with $ux = 1$. Then u^d exists if and only if u^{-1} exists.*

PROOF.   This follows immediately from the fact that

$$i(u) = r(u) = l(u) = 0.$$

LEMMA 2.   *Let R be a ring with unity 1 and let $a \in R$. Then a^* exists if and only if a^2 and a are associates. In either case a is unit regular and a^2 is equivalent to a.*

PROOF.   Suppose a^* exists. Then $u = a^* + (1 - aa^*)$ is *invertible* with inverse $u^{-1} = a + (1 - aa^*)$. Hence

$$au = aa^{\#} = ua, \quad aua = a \quad \text{and} \quad a^2u = a = ua^2.$$

The converse is clear.

The local result of this lemma should be compared with the results obtained in Ehrlich (1968), Hartwig (submitted). It is well-known that in the case of matrices over a field, the equivalence between A and A^2 is also sufficient for $A^{\#}$ to exist. This is also true for a skew field as shown in Section 5. We note the following global analogue of the above result, namely that a regular ring with unity is strongly regular (that is $l(a) \leqq 1$ for all $a \in R$) if and only if every element in R is equivalent to its square. Indeed, the necessity is contained in Lemma 2, while conversely, if $p_k a^{2k} q_k = a^k$, $k = 1, 2, \cdots$ and if $a^n = 0$, $n \geqq 2$, then $a^{n-1} = p_{n-1} a^{2n-2} q_{n-1} = 0$. Consequently R has no nonzero nilpotent elements and thus must be strongly regular since all idempotents are central. It is an open question whether the regularity assumption can be dropped. Let us now turn to the question of existence of the Group and Drazin inverses of triangular matrices. We shall see that the existence of the group inverse requires strong additional consistency conditions.

2. Drazin inverses of triangular matrices

THEOREM 1. *Let R be a ring with unity 1 and let $M = \begin{bmatrix} u & 0 \\ v & w \end{bmatrix} \in R_{2 \times 2}$,*

with $M^l = \begin{bmatrix} u^l & 0 \\ y_l & w^l \end{bmatrix}$ $l = 1, 2, \cdots$. If $u^d[w^d]$ exists, then

(i)   *M^d exists if and only if $w^d[u^d]$ exists. In which case*

$$M^d = \begin{bmatrix} u^d & 0 \\ b & w^d \end{bmatrix} \quad \text{and} \quad MM^d = \begin{bmatrix} uu^d & 0 \\ fy_k(u^d)^k + (w^d)^k y_k e & ww^d \end{bmatrix}$$

where $e = 1 - uu^d$, $f = 1 - ww^d$,

(2.0) $b = -w^d v u^d + fy_k(u^d)^{k+1} + (w^d)^{k+1} y_k e,$

         $k \geqq i(M)$ *and* $\max\{i(u), i(w)\} \leqq i(M) \leqq i(u) + i(w).$

(ii)   *$M^{\#}$ exists if and only if $u^{\#}$ and $w^{\#}$ exist and*

$$(1 - ww^{\#})v(1 - uu^{\#}) = 0.$$

In which case

$$M^{\#} = \begin{bmatrix} u^{\#} & 0 \\ b & w^{\#} \end{bmatrix}, \qquad MM^{\#} = \begin{bmatrix} uu^{\#} & 0 \\ fvu^{\#} + w^{\#}ve & ww^{\#} \end{bmatrix}$$

where $e = 1 - uu^{\#}$, $f = 1 - ww^{\#}$, and

$$b = -w^{\#}vu^{\#} + fv(u^{\#})^2 + (w^{\#})^2 ve.$$

PROOF.   (i) Suppose first that u^d and M^d exist, with $l = i(M)$, and let $M^d = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$. Then $M^{l+1}M^d = M^l$ shows that

(2.1) $u^{l+1}a = u^l,$

(2.2) $u^{l+1}c = 0,$

(2.3) $y_{l+1}a + w^{l+1}b = y_l.$

Moreover, $M^dM^{l+1} = M^l$ implies by symmetry that

(2.4) $w^l = dw^{l+1}$

(2.5) $bu^{l+1} + dy_{l+1} = y_l,$

while the identity $M^d = M^dMM^d$ yields:

(2.6) $b = bua + d(va + wb).$

Lastly, the identities $M^d = (M^d)^{l+1}M^l = M^l(M^d)^{l+1}$ give:

(2.7) $c = u^l x = x_1 w^l$ for some x, x_1 in R

and

(2.8) $a = u^l y,\quad d = y_1 w^l$ for some y, y_1 in $R.$

Note that (2.1) implies that $i(u) \leqq i(M)$. Hence, combining (2.2) and (2.7) yields

(2.9) $c = u^l x = (uu^d)^{l+1}u^l x = (u^d)^{l+1}u^{l+1}c = 0,$

while combining (2.1) and (2.8) shows that

(2.10) $a = a^l y = (uu^d)^{l+1}u^l y = (u^d)^{l+1}u^{l+1}a = u^d.$

Similarly (2.4) and (2.8) combine to show that $d = w^d$ and $i(w) \leqq i(M)$. Rewriting (2.6) we arrive at

(2.11) $b = w^d vu^d + buu^d + ww^d b.$

Next, observe that $y_k = w^{k-1}v + w^{k-2}vu + \cdots + vu^{k-1}$, which implies that

(2.12) $y_{l+1} = w^l v + y_l u = vu^l + wy_l.$

Premultiplication of (2.3) by $(w^d)^{l+1}$ now yields

$$ww^d b = (w^d)^{l+1}w^{l+1}b = (w^d)^{l+1}(y_l - y_{l+1}u^d)$$

(2.13) $$= (w^d)^{l+1}[y_l - (y_l u + w^l v)u^d]$$

$$= (w^d)^{l+1}y_l(1 - uu^d) - w^d vu^d.$$

Similarly, post multiplication of (2.5) by $(u^d)^{l+1}$ gives:

(2.14) $$buu^d = (1 - ww^d)y_l(u^d)^{l+1} - w^d vu^d.$$

Lastly, substituting the latter two expressions in (2.11) we arrive at the desired expression for b.

Conversely, suppose that u^d and w^d exist and consider the matrix $X = \begin{bmatrix} u^d & 0 \\ b & w^d \end{bmatrix}$, where b is given as in (2.0), with k an integer to be determined. We shall now verify that for $k \geq i(u) + i(w)$, X is indeed the Drazin inverse of M. First of all, by (2.11), $XMX = X$ is satisfied for all k. Next, the commutativity condition $MX - XM = 0$ requires that

(2.15) $$wb - bu - w^d v + vu^d = 0.$$

Substituting for b and using the following identities

(2.16)(a) $$y_k(u^d)^k - wy_k(u^d)^{k+1} = vu^d - w^k v(u^d)^{k+1}$$

(2.16)(b) $$(w^d)^k y_k - (w^d)^{k+1} y_k u = w^d v - (w^d)^{k+1} vu^k$$

we obtain in (2.15)

$$(1 - ww^d)w^k v(u^d)^{k+1} - (w^d)^{k+1}vu^k(1 - uu^d),$$

which vanishes if we take $k \geq i(u)$, $k \geq i(w)$. Lastly, the equations $M^{k+1}X = M^k = XM^{k+1}$ will be satisfied provided

(2.17) $$(1 - ww^d)y_k(1 - uu^d) = 0.$$

Indeed, $w^{k+1}b = ww^d(y_k - y_{k+1}u^d)$ becomes with aid of (2.12) and (2.0),

$$- w^k vu^d + ww^d y_k(1 - uu^d) = y_k(1 - uu^d) - w^k vu^d,$$

which reduces to (2.17). It is now clear from the expression for y_k, that the consistency condition (2.17) will be satisfied if k is taken *large enough*. In particular $k = i(u) + i(w)$ would suffice. Hence by the uniqueness of the solution to (1.2) $X = M^d$ and

(2.18) $$\max\{i(u), i(w)\} \leq i(M) \leq i(u) + i(w).$$

Thus for example, if w is invertible then $i(w) = 0$ and $i(M) = i(u)$! Using (2.16), the required form for MM^d is now easily obtained, completing the proof or part (i).

(ii) Suppose first that u^d and M^* exist with $M^* = \begin{bmatrix} u^d & 0 \\ b & w^d \end{bmatrix}$. Since by part (i), $\max\{i(u), i(w)\} \leq 1$ we have $u^d = u^*$, $w^d = w^*$. By (2.17) the equation $M^2 X = M$ will be satisfied only if $(1 - ww^*)v(1 - uu^*) = 0$, since $y_1 = v$. The converse follows immediately on verifying that the expression for

M^{*} does obey (1.1), which has a unique solution, namely the group inverse of M.

Before turning to the immediate corollaries of this theorem, a few remarks should be made here.

REMARKS. 1. In a general ring the existence of a solution to the equation $u^{l+1}a = u^{l}$, does *not* ensure the existence of u^{d}. In fact even if M is invertible one cannot conclude this, as seen from the following example:

$$\begin{bmatrix} u & 0 \\ 1 & w \end{bmatrix}\begin{bmatrix} w & 1-wu \\ -1 & u \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} w & 1-wu \\ -1 & u \end{bmatrix}\begin{bmatrix} u & 0 \\ 1 & w \end{bmatrix}$$

with $uw = 1 \neq wu$.

2. It follows from Theorem 1, that strong regularity is *not* inherited by $R_{2\times2}$. It is an open question whether this is true for strong π-regularity, that is, when every element has a Drazin inverse.

3. For real matrices the expression for M^{*} was also stated in Meyer (1975).

4. In Theorem 1, one may obviously replace u, v and w by *conformable matrices* over R, since all identities used in the proof then remain valid.

5. The matrix $\begin{bmatrix} b & w^{*} \\ u^{*} & 0 \end{bmatrix}$ is a 1-2 inverse, (Ben-Israel and Greville (1974), page 7) for the matrix $\begin{bmatrix} 0 & u \\ w & v \end{bmatrix}$, but does *not* commute with the latter. In fact it is easily seen that the group inverse of $\begin{bmatrix} 0 & u \\ w & v \end{bmatrix}$, if any, cannot be triangular unless the matrix is invertible. We shall return to this in Section 4.

6. If M^{*} and u^{*} exist, then by (2.5) with $k = 1$, $w[(w^{*})^{2}(wv + vu)] + (bu)u = v$. Using Roth's factorization, (Roth (1952)), we have

$$\begin{bmatrix} 1 & 0 \\ -x & 1 \end{bmatrix}\begin{bmatrix} u & 0 \\ v & w \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -y & 0 \end{bmatrix} = \begin{bmatrix} u & 0 \\ 0 & w \end{bmatrix} = D,$$

where $x = bu$ and $y = w^{*}v + (w^{*})^{2}vu$. Thus M^{*} is equivalent to a diagonal matrix D. Similarly,

$$\begin{bmatrix} 1 & 0 \\ -p & 1 \end{bmatrix}\begin{bmatrix} u^{2} & 0 \\ wv + vw & w^{2} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -q & 1 \end{bmatrix} = \begin{bmatrix} u^{2} & 0 \\ 0 & w^{2} \end{bmatrix},$$

where $p = wb + bu$ and $q = (w^{*})^{2}(wv + vu) + (w^{*})^{3}(wv + vu)u$. As expected, this shows with aid of Lemma 2 that M^{2} is equivalent to M, as D^{2} is equivalent to D.

7. The inequality (2.18) is best possible since it cannot be improved for

matrices over a field. Indeed if $M = \begin{bmatrix} U & 0 \\ V & W \end{bmatrix}$, where $U \in \mathscr{F}_{m \times m}$ and $W \in \mathscr{F}_{n \times n}$, and if ψ_M denotes the minimal polynomial of M then it is well-known that

$$\text{LCM}(\psi_U, \psi_V) \mid \psi_M \mid \psi_U \cdot \psi_V.$$

There is a universal choice of v for which the left hand inequality in (2.18) is attained, for all choices of u and v, namely $v = 0$. It is unknown for what choices of v the right hand inequality is attained universally.

8.　If $\begin{bmatrix} u & 0 \\ v & w \end{bmatrix}$ has a Drazin inverse, so has $\begin{bmatrix} w & v \\ 0 & u \end{bmatrix}$. Indeed, it was shown by Cline (1965) that $(ab)^d$ exists if and only if $(ba)^d$ exists, in which case they are related by the equation

(2.19) $$(ab)^d = a[(ba)^d]^2 b.$$

It follows further from the proof that the indices $i(ab)$ and $i(ba)$ differ *at most* by unity. This should be compared with the corresponding results for matrices over a field, for which the minimal polynomials of AB and BA differ at most by a factor of λ. Thus if $(ba)^{-1}$ exists then $i(ab) = 0$ or 1 so that ab must have a group inverse. On the other hand if $(ba)^*$ exists then $i(ab) = 0, 1$ or 2. In this case $(ab)^d = (ab)^* \Leftrightarrow a[1 - (ba)(ba)^*]b = 0$ while $(ab)^d = (ab)^{-1} \Leftrightarrow a(ba)^* b = 1$. In particular if $(ba)^*$ exists and $xb = 1 = ay$ then the identity $xba(ba)^* bay = xbay$ shows that $a(ba)^* b = 1$ guaranteeing that ab has an inverse. This generalizes the results of Theorem 2 p. 163 in (Ben-Israel and Greville (1974)).

COROLLARY 1.　*The matrix* $M = \begin{bmatrix} 0 & 0 \\ v & w \end{bmatrix}$ *has a Drazin inverse if and only if* w^d *exists. In which case* $M^d = \begin{bmatrix} 0 & 0 \\ (w^d)^2 v & w^d \end{bmatrix}$, *with*

$$i(M) = \begin{cases} 1 & \text{if} \quad i(w) = 0 \\ i(w) & \text{if} \quad (1 - ww^d)w^{i(w)-1}v = 0 \\ i(w) + 1 & \text{if} \quad (1 - ww^d)w^{i(w)-1}v \neq 0. \end{cases}$$

PROOF.　Observe that $M^{k+1}M^d = \begin{bmatrix} 0 & 0 \\ w^k w^d v & w^k \end{bmatrix}$, while $M^{k+2}M^d = M^{k+1}$ if $k \geq i(w)$. Note that this implies that M^* exists *exactly* when w^* exists and $v = ww^*v$.

COROLLARY 2.　*Let* $M = \begin{bmatrix} u & 0 \\ v & u \end{bmatrix} \in R_{2 \times 2}$. *Then*

(i) M^d exists if and only if u^d exists. In which case

$$M^d = \begin{bmatrix} u^d & 0 \\ b & u^d \end{bmatrix}, \qquad MM^d = \begin{bmatrix} uu^d & 0 \\ ey_k(u^d)^k + (u^d)^k y_k e & uu^d \end{bmatrix}.$$

where $e = 1 - uu^d$, $b = -u^d vu^d + ey_k(u^d)^{k+1} + (u^d)^{k+1} y_k e$, $y_k = u^{k-1}v + u^{k-2}vu + \cdots + vu^{k-1}$, $k \geq i(M)$ and $i(u) \leq i(M) \leq 2i(u)$.

(ii) M^* exists if and only if u^* exists and $(1 - uu^*)v(1 - uu^*) = 0$. In which case

$$M^* = \begin{bmatrix} u^* & 0 \\ -u^*vu^* + u^{*2}ve + evu^{*2} & u^* \end{bmatrix} \qquad MM^* = \begin{bmatrix} uu^* & 0 \\ evu^* + u^*ve & uu^* \end{bmatrix}$$

where $e = 1 - uu^*$.

PROOF.   Observe that (2.1) and (2.4) now *ensure* that u^d exists (unlike in the case where $u \neq w$!) and $i(u) \leq k$. The rest is immediate from Theorem 1.

COROLLARY 3.   *Let*

$$M = \begin{bmatrix} u_1 & & & 0 \\ & u_2 & & \\ & & \cdot & \\ ? & & & \cdot \\ & & & & u_n \end{bmatrix} \in R_{n \times n}.$$

(i)   *If M and all but one of the u_i have Drazin inverses (inverses, group inverses) then all u_i possess the same type of inverse. The converse is also true in the first two cases. In which case* $i(M) \leq \Sigma_{k=1}^n i(u_k)$.

(ii)   *If all $u_i = u$, then*

$$M^d, (M^{-1}, M^*) \text{ exists } \Rightarrow u^d(u^{-1}, u^*) \text{ exists.}$$

The converse is also true in the first two cases.

PROOF.   (i) If u_i^d exists for all i, partition M as

(2.20) $$M = \begin{bmatrix} u_1 & & & & \vline & \\ & u_2 & & & \vline & \\ & & \cdot & & \vline & 0 \\ ? & & \cdot & & \vline & \\ & & & u_k & \vline & \\ \hline & & & & \vline & u_{k+1} & \vline 0 \\ & ? & & & \vline & & \cdot \\ & & & & \vline & & & u_n \end{bmatrix} \quad k = 1, 2, \cdots, n-1,$$

and apply Theorem 1(i) for each value of k, ensuring that M^d exists. Conversely, suppose that M^d and *all u_i^d* exist, *except* possibly u_j^d. Then we partition M as:

$$(2.21) \qquad\qquad M = \begin{bmatrix} U & 0 & 0 \\ ? & u_j & 0 \\ ? & ? & W \end{bmatrix},$$

in which U^d and W^d exist as shown above. Hence by Theorem 1, and Remark 4, $\begin{bmatrix} U & 0 \\ ? & u_j \end{bmatrix}$ exists, which guarantees that u_j^d does also exist. Suppose next, that $M^\ast$ and all $u_i^\ast$, $i \neq j$, exist. Again partitioning M as in (2.20) for $k = 1, 2, \cdots, j - 1$, shows with aid of Theorem 1(ii) that $\left[\dfrac{u_j}{?} \middle| \dfrac{0}{W} \right]^\ast$ exists. On the other hand, using (2.21) shows then that $u^\ast$ exists. Similarly $W^\ast$ exists and hence again by Theorem 1, $u_j^\ast$ exists.

The remaining case where the inverses exist is proven similarly.

(ii)   This follows from part (i) and Corollary 2.

Let us now turn to the Drazin inverses of triangular Toeplitz and bi-diagonal matrices.

3. The Drazin inverse of a triangular Toeplitz matrix

THEOREM 2.   *Let R be a ring with unity and let*

$$M_n = \begin{bmatrix} a_0 & & & 0 \\ a_1 & \cdot & & \\ \vdots & \cdot & \cdot & \\ \cdot & & \cdot & \\ a_{n-1} & & a_1 & a_0 \end{bmatrix} = a_0 I + a_1 J + a_2 J^2 + \cdots + a_{n-1} J^{n-1},$$

be a triangular Toeplitz matrix, where $J = [\delta_{i,j+1}]_1^n$. Then

(i)   *M_n^d exists if and only if a_0^d exists. In which case M_n^d is also triangular Toeplitz*

$$M_n^d = b_0 I + b_1 J + \cdots + b_{n-1} J^{n-1}, \quad where$$

$$b_0 = a_0^d$$

$$b_1 = -b_0 a_1 b_0 + e y_1^{(k)} (b_0)^{k+1} + b_0^{k+1} y_1^{(k)} e$$

$$(3.1) \qquad \vdots$$

$$b_j = -\sum_{i=1}^{j} b_{j-i} a_i b_0 + y_j^{(k)} b_0^{k+1} - \sum_{i=1}^{j} c_{j-i} y_i^{(k)} b_0^{k+1}$$

$$+ \sum_{i=1}^{j} z_{j-i}^{(k)} y_i^{(k)} e,$$

in which $e = 1 - b_0 a_0$, $k \geq i(M)$ and c_j, $y_j^{(k)}$, $z_j^{(k)}$ are coefficients of J^j in $M_n M_n^\ast$, M_n^k and $(M_n^d)^k$ respectively.

(ii) $M_n^\#$ *exists if and only if* $a_0^\#$ *exists and*

(3.2) $CC_k:$ $eX_ke = 0$ $k = 0, 1, \cdots, n-2,$

where

$$X_k = p_ka_1 + p_{k-2}a_2 + \cdots + p_0a_{k+1}, \qquad p_0 = 1$$

and

(3.3) $$p_k = \sum_{\substack{i_1+i_2+\cdots+i_k = k \\ \text{no repetition}}} (-1)^k a_{i_1} a_0^\# a_{i_2} a_0^\# \cdots a_{i_k} a_0^\#.$$

In which case

$$b_0 = a_0^\#$$

$$b_1 = -a_0^\# a_1 a_0^\# + ea_1 a_0^{\#2} + (a_0^\#)^2 a_1 e$$

(3.4)

$$b_k = -(b_{n-1}a_1 + \cdots + b_0a_k)b_0 - (c_{k-1}a_1 + \cdots + c_0a_k)b_0^2$$

$$+ a_k b_0^2 + \sum_{i=1}^{k} \sum_{j=1}^{i} b_{k-i}b_{i-j}a_je, \qquad k = 2, 3, \cdots, n-1.$$

PROOF.   (i) The first part was shown in Corollary 3. Suppose therefore that M_n^d and a_0^d exist, and that $i(M_n) = l.$ We shall first show that M_n^d *must* also be triangular Toeplitz.

For each $k = 1, 2, \cdots, n$ and $r = 1, 2, \cdots$ we partition $(M_n)'$ and M_n^d conformally as

(3.5) $$M_n' = \begin{bmatrix} M_k^r & 0 \\ ? & M_{n-1}^r \end{bmatrix} \quad \text{and} \quad M_n^d = \begin{bmatrix} A_k & C_k \\ B_k & D_k \end{bmatrix},$$

where $M_1 = a_0$ and M_0 is absent. Comparing block components in

$$M_n^{l+1} M_n^d = M_n^l = M_n^d M_n^{l+1},$$

we find that

$$M_k^{l+1} A_k = M_k^l = D_{r-k} M_n^{l+1},$$

for each $k = 1, 2, \cdots, n,$ so that M_k^d exists for these values of k. Moreover, by Theorem 1, $A_k = M_k^d$ and $C_k = 0$ for $k = 1, 2, \cdots, n.$ Thus M_n^d is lower triangular. If we now partition

$$M_n = \begin{bmatrix} M_i & 0 & 0 \\ ? & M_j & 0 \\ ? & ? & M_k \end{bmatrix}, \quad M_n^d = \begin{bmatrix} M_i^d & 0 & 0 \\ ? & X & 0 \\ ? & ? & M_k^d \end{bmatrix},$$

for all $i + j + k = n$, then it is easily seen that $X = M_j^d$. On taking in succession $j = 2, 3, \cdots, n$, for each value of j, taking $i = 0, 1, 2, \cdots$, one sees that M_n^d must be striped, that is, have the same elements along any diagonal, say $M_n^d = b_0 I + b_1 J + \cdots + b_{n-1} J^{n-1}$, with $b_0 = a_0^d$. It is now clear that M_n^k, $(M_n^d)^k$ and $M_n M_n^d$ are all triangular Toeplitz, say

$$M_n^k = \sum_{j=0}^{n-1} y_j^{(k)} J^j, \qquad (M_n^d)^k = \sum_{j=0}^{n-1} z_j^{(k)} J^j \quad \text{and}$$

(3.6)

$$M_n M_n^d = \sum_{j=0}^{n-1} c_j J^j.$$

Partitioning M_n, M_n^d and M_n^k as

$$M_n = \left[\begin{array}{c|c} a_0 & 0 \\ \hline \underline{a} & M_{n-1} \end{array} \right], \quad M_n^d = \left[\begin{array}{c|c} b_0 & 0 \\ \hline \underline{b} & M_{n-1}^d \end{array} \right], \quad M_n^k = \left[\begin{array}{c|c} a_0^k & 0 \\ \hline \underline{y}^{(k)} & M_{n-1}^k \end{array} \right],$$

where $\underline{a}^T = [a_1, \cdots, a_{n-1}]$, $\underline{b}^T = [b_1, \cdots, b_{n-1}]$ and $\underline{y}^{(k)T} = [y_1^{(k)}, \cdots, y_{n-1}^{(k)}]$, then we obtain form (2.0);

(3.7) $\underline{b} = -M_{n-1}^d \underline{a} b_0 + (I - M_{n-1} M_{n-1}^d) \underline{y}^{(k)} b_0^{k+1} + (M_{n-1}^d)^{k+1} \underline{y}^{(k)} e,$

or

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \end{bmatrix} = -M_{n-1}^d \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n-1} \end{bmatrix} b_0 + (I - M_{n-1} M_{n-1}^d) \begin{bmatrix} y_1^{(k)} \\ y_2^{(k)} \\ \vdots \\ y_{n-1}^{(k)} \end{bmatrix} b_0^{k+1} + M_{n-1}^d \begin{bmatrix} y_1^{(k)} \\ y_2^{(k)} \\ \vdots \\ y_{n-1}^{(k)} \end{bmatrix} e,$$

where $e = 1 - b_0 a_0$. Equating entries in (3.7) now yields (3.1).

(ii) If $M_n^\#$ exists, then by Corollary 3, $a_0^\#$ exists. Moreover from Theorem 1 (ii), the consistency conditions

(3.8) $(I - M_{n-1} M_{n-1}^\#) \underline{a} e = \underline{0},$

must hold. (This says that $\underline{a} e$ must be an eigenvector of $M_{n-1} M_{n-1}^\#$ associated with eigenvalue 1.) In addition, (3.7) now becomes:

(3.9) $\underline{b} = -M_{n-1}^\# \underline{a} b_0 + (I - M_{n-1} M_{n-1}^\#) \underline{a} b_0^2 + (M_{n-1}^\#)^2 \underline{a} e,$

where $b_0 = a_0^\#$ and $e = 1 - b_0 a_0$. On equating entries we obtain the quadratic recurrence relation (3.4), from which we may generate b_k recursively.

Let us now turn to the problem of implicitly eliminating the b_k's from the consistency conditions (3.8), which may conveniently be written as

(3.10) $(c_k a_1 + c_{k-1} a_2 + \cdots + c_1 a_k) e = e a_{k+1} e, \qquad k = 0, 1, \cdots, n - 2,$

in particular $e a_1 e = 0$.

We begin by observing the following identities which are really a consequence of the triangular Toeplitz structure of our matrices.

$$\text{(i)} \quad M_{n-1}\underline{c} + \underline{a}c_0 = \underline{a} \quad \text{or} \quad M_{n-1}\underline{c} = \underline{a}e$$

$$\text{(ii)} \quad M_{n-1}\underline{b} + \underline{a}b_0 = \underline{c} = M^{\#}_{n-1}\underline{a} + \underline{b}a_0$$

(3.11) $\quad$ $$\text{(iii)} \quad M^{\#}_{n-1}M_{n-1}\underline{c} = \underline{c}e$$

$$\text{(iv)} \quad M_{n-1}M^{\#}_{n-1}\underline{a} + \underline{c}a_0 = \underline{a} \quad \text{or} \quad (I - M_{n-1}M^{\#}_{n-1})\underline{a} = \underline{c}a_0$$

$$\text{(v)} \quad M^{\#}_{n-1}\underline{c} + \underline{b}c_0 = \underline{b} \quad \text{or} \quad M^{\#}_{n-1}\underline{c} = \underline{b}e,$$

where $\underline{c}^T = [c_1, c_2, \cdots, c_{n-1}]$. Using the second of these in (3.9), multiplied by M_{n-1}, we obtain

$$(3.12) \qquad \underline{c} = M_{n-1}\underline{b} + \underline{a}b_0 = (I - M_{n-1}M^{\#}_{n-1})\underline{a}b_0 + M^{\#}_{n-1}\underline{a}e,$$

which, on using the second and fourth identities reduces to the triviality $\underline{c} = \underline{c}(1 - e) + \underline{c}e$. We may still equate components in (3.12), however, to obtain

$$c_k = -(c_{k-1}a_1 + c_{k-2}a_2 + \cdots + c_1a_{k-1})b_0 + ea_kb_0 + c_ke,$$
(3.13)
$$k = 1, 2, \cdots, n - 1,$$

where $c_1 = ea_1b_0 + b_0a_1e$.

We now use the following result which is proven by induction:

LEMMA 3. *If* $y_k = (y_{k-1}\alpha_1 + y_{k-2}\alpha_2 + \cdots + y_1\alpha_{k-1}) + \beta_{k-1} \; k \geq 2$ *then*

$$y_k = y_1p_{k-1} + q_{k-1}, \quad k \geq 1,$$

where

$$p_k = \sum_{\substack{i_1 + \cdots + i_k = k \\ \text{no repetition}}} \alpha_{i_1} \cdots \alpha_{i_k} \equiv p_{k-1}\alpha_1 + p_{k-2}\alpha_2 + \cdots + p_0\alpha_k,$$

$$(3.14) \qquad q_k = \beta_1 p_{k-1} + \beta_2 p_{k-2} + \cdots + \beta_k p_0, \quad \text{and} \quad p_0 = 1, \; q_0 = 0.$$

For example, $p_1 = \alpha_1$, $p_2 = \alpha_1^2 + \alpha_2$, $p_3 = \alpha_1^3 + \alpha_2\alpha_1 + \alpha_1\alpha_2 + \alpha_3$ etc. We remark before continuing, that p_k should not be confused with the coefficient of x^k in $(1 + \alpha_1x + \alpha_2x^2 + \cdots)^k$, in which repetition *does* occur, and that also

$$(3.15) \qquad p_k = \alpha_1p_{k-1} + \alpha_2p_{k-2} + \cdots + \alpha_kp_0.$$

Let us now apply Lemma 3 to (3.13) with $\alpha_i = -a_ib_0$, $\beta_i = ea_{i+1}b_0 + c_{i+1}e$. Then:

$$(3.16) \qquad c_k = c_1p_{k-1} + q_{k-1},$$

where p_k is given as in (3.3) and q_k as in (3.14). Substituting (3.16) in the consistency conditions (3.10) and using $c_1 = ea_1b_0 + b_0a_1e$, we arrive at

$$ea_ke = c_1 \sum_{i=1}^{k-1} p_{k-i-1}a_ie + \sum_{i=1}^{k-1} q_{k-i-1}a_ie,$$

(3.17)

$$= \sum_{i=1}^{k-1} (ea_1b_0p_{k-i-1} + q_{k-i-1})a_ie + b_0a_1e \sum_{i=1}^{k-1} p_{k-i-1}a_ie.$$

Substituting for q_{k-i-1} from (3.14) and recalling the definition of β_i shows that

$$ea_ke = \sum_{i=1}^{k-1} \left[ea_1b_0p_{k-i-1} + \sum_{j=1}^{k-i-1} (ea_{j+1}b_0 + c_{j+1}e)p_{k-i-j-1} \right] a_ie$$

$$+ b_0a_1eX_{k-2}e,$$

where X_k is given as in (3.2). Now using (3.15) we obtain

$$ea_ke = -\sum_{i=1}^{k-1} ep_{k-i}a_ie + \sum_{i=1}^{k-1} \sum_{j=1}^{k-i-1} c_{j+1}ep_{k-i-j-1}a_ie$$

(3.18)

$$+ b_0a_1eX_{k-2}e.$$

Interchanging summations in the latter term we obtain

$$\sum_{j=1}^{k-2} c_{j+1} \sum_{i=1}^{k-j-1} ep_{k-i-j-1}a_ie = \sum_{j=1}^{k-2} c_{j+1}eX_{k-j-2}e,$$

so that the consistency condition (3.18) collapses to

(3.19) $0 = eX_{k-1}e + b_0a_1eX_{k-2}e + c_2eX_{k-3}e + \cdots + c_{k-1}eX_0e.$

$$k = 1, 2, \cdots, n - 1.$$

Since the *first* consistency condition is indeed equal to $eX_0e = 0$, it is now clear that (3.10) is equivalent to the conditions $eX_ie = 0$, $i = 0, 1, \cdots, n - 2$. Conversely, if $a_0^\#$ exists, and $eX_ie = 0$ $i = 0, \cdots, n - 2$, let

$$B_n = b_0I + b_1J + \cdots + b_{n-1}J^{n-1},$$

where the b_i are generated recursively by (3.4). Then by Theorem 1 (ii) and the equivalence of (3.2) with (3.8) it follows that for each $k = 1, 2, \cdots, n$, $B_k = M_k^\#$, completing the proof.

Of special interest is the bidiagonal case:

COROLLARY 4.   *If $M_n = uI + vJ \in R_{n \times n}$, where $J = [\delta_{i,j+1}]_1^n$, then $M^\#$ exists if and only if $u^\#$ exists, and $e(vu^\#)^ke = 0$ for $k = 0, 1, \cdots, n - 2$, where $e = 1 - uu^\#$. In which case*

$$M_n^\# = b_0I + b_1J + \cdots + b_{n-1}J^{n-1},$$

where

$$b_0 = u^{\#}, \quad b_1 = -u^{\#}vu^{\#} + evu^{\#2} + u^{\#2}ve$$

and

$$
\begin{aligned}
b_k = &-b_{k-1}vu^{\#} - (ub_{k-1} + vb_{k-2})vu^{\#2} \\
(3.20) \quad &+ (b_{k-1}b_0 + b_{k-2}b_1 + \cdots + b_0b_{k-1})ve \quad \text{for} \quad k = 2, 3, \cdots, n-1.
\end{aligned}
$$

Before turning to the Drazin inverses of companion and hyper-companion matrices, several remarks are appropriate here.

REMARKS. 1. As expected, the consistency conditions (3.2) for $M^{\#}$ to exist involve *all* entries a_i of M.

2. The process of equating coefficients in (3.7), (3.8) and (3.9) may be facilitated on writing $M_n = a_0I + \cdots + a_{n-1}J^{n-1}$, $M_n^d = b_0I + b_1J + \cdots$ etc. and equating powers of J. It does not, however, seem possible to use these representations *directly* (without using blocks) to prove that, for example, M_n^d exists if and only if a_0^d exists. Indeed we have seen that if N is any *lower* matrix, that is $n_{ij} = 0$ for $i \leq j$, then $(uI + N)^d$ exists exactly when U^d exists.

We conjecture that this is true for all nilpotent matrices N for which $\Sigma_{i=1}^k r_iNs_i$ is nilpotent for all $r_i, s_i \in R$ and all $k = 1, 2, \cdots$. When $uN = Nu$ we may apply the following:

LEMMA 4. *Let* $A, N \in R_{n \times n}$ *such that* $N^k = 0$, *and* $AN = NA$. *Then* A^d *exists in* $R_{n \times n}$ *if and only if* $(A + N)^d$ *exists in* $R_{n \times n}$, *in which case*

$$(A + N)^d = A^d(I + A^dN)^{-1}, \quad \text{and}$$

$$A^d = (A + N)^d(I - (A + N)^dN)^{-1}.$$

The commutivity may not be dropped in this, as seen from the 4×4 integer matrix:

$$
M = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = E + N \in Z_{4 \times 4}
$$

which has a unique Drazin inverse in $Q_{4 \times 4}$, but not in $Z_{4 \times 4}$. Even if $AB = BA$ and A^d, B^d exist in $R_{n \times n}$, $(A + B)^d$ may not exist in $R_{n \times n}$, as seen from the example where $A = B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \in Z_{2 \times 2}$.

3.   If u is a unit, then as expected no consistency conditions are necessary as M^{-1} also exists and is also striped! c.f. (Huang and Cline (1972)). If $u = v$, then $M = u(I + J)$ and $ev = 0$, so that again the consistency conditions hold automatically and $M^\# = u^\#(I - J + J^2 - \cdots \pm J^{n-1})$. If $v = 1$, then the identity $eve = e^2 = e = 0$ shows that $M^\#$ cannot exist unless u is invertible. Moreover if u is nilpotent of index k, then M is nilpotent of index $n + k - 1$.

4.   The identity $M^{\#2}M = M^\#$ yields the identity

$$u(b_0 b_k + b_1 b_{k-1} + \cdots + b_k b_0) + v(b_0 b_{k-1} + \cdots + b_{k-1} b_0) = b_k \quad k = 0, 1, \cdots$$

which should be helpful in finding a closed form expression for b_k.

5.   Even in the bidiagonal case $M_n = uI + vJ$, no closed form formula is known at the present for the entries b_k in M_n^d or $M_n^\#$. The main difficulty coming from the *non-linearity* of the recurrence relation (3.20). It may, however, be shown by induction that when $M^\#$ exists, b_k has the form

$$b_k = (-1)^k (u^\# v)^k u^\# + \sum_{i=0}^{k-1} [(u^\# v)^i e\alpha_i^{(k)} + \beta_i^{(k)} e(v^\# u)^i]$$

for some suitable coefficients $\alpha_i^{(k)}$, $\beta_i^{(k)}$. For example,

$$b_2 = (u^\# v)^2 u^\# - evu^{\#2} vu^\# - evu^\# vu^{\#2} - u^\# vu^{\#2} ve - u^{\#2} vu^\# ve$$
$$- u^{\#2} vevu^\# - u^\# vevu^{\#2} + evu^{\#3} ve.$$

6.   If the consistency conditions $eX_i e = 0$ hold for $i = 0, 1, \cdots, k - 1$, but $eX_k e \neq 0$, then $i(M_{k+1}) \leq 1$ and hence $i(M_n) \leq n - k$. It would be desirable to find a simpler proof of the consistency condition (3.2) using, say, graph theory.

4. Group inverses and Drazin inverses of companion matrices

Throughout this section we denote the lower companion matrix, the hypercompanion matrix and the Hankel matrix associated with the monic polynomial $p(\lambda) = p_0 + p_1 \lambda + \cdots + \lambda^n$ by

(4.1)

$$L[p(\lambda)] = \begin{bmatrix} 0 & & & & & -p_0 \\ 1 & & & & & -p_1 \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & \ddots & \\ 0 & & & & 1 & -p_{n-1} \end{bmatrix}, \quad H[p^m(\lambda)] = \begin{bmatrix} L(p) & 0 & & & \\ N & L(p) & & & \\ & N & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & N & L(\end{bmatrix}$$

and

$$G[p(\lambda)] = \begin{bmatrix} p_1 & p_2 & \cdot & \cdot & \cdot & \cdot & p_n \\ p_2 & & & & \cdot & \cdot & \\ \cdot & \cdot & & & & & \\ \cdot & & & & & & \\ \cdot & & & & & & \\ p_n & \cdot & & & & & \end{bmatrix}$$

respectively, where N is the basis matrix $E_{1,n}$. The appearance of companion matrices in very general settings is essentially due to the basic cyclic decomposition theorems for modules over suitable rings (Jacobson (1943), Chapter 3).

We shall first give several results for arbitrary rings illustrating the "cyclic nature" of L, thereby generalizing some well-known results for fields. These are useful in the theory of *block* companion matrices, with $R = \mathscr{F}_{n \times n}$, $\mathscr{F}$ a field, which appear naturally in the study of systems of differential and difference equations (Lancaster (1969), p. 193). The left and right functional values of the polynomial $p(\lambda)$ evaluated at the matrix X (Gantmacher (1960), p. 81), will be denoted by $p_L(X) = \Sigma_{i=0}^n X^i p_i$ and $p_R(X) = \Sigma_0^n p_i X^i$ respectively, in which Xp_i and p_iX indicate entrywise multiplication. A crucial difference between the commutative and non-commutative case is that the evaluation map $\Gamma_A : p(\lambda) \to p_L(A)$ does not preserve products, in that $\Gamma_A[p(\lambda)q(\lambda)] \neq p_L(A)q_L(A)$. As always, the unit vectors $[0, 0 \cdots, 1, \cdots 0]^T$, will be denoted by e_i.

THEOREM 3.   *Let R be a ring with unity 1; let $R_{n \times n}(\lambda)$ denote the ring of λ-matrices over R and suppose L, H and G are the lower companion, the hypercompanion and the Hankel matrices associated with $p(\lambda) = p_0 + p_1\lambda + \cdots + \lambda^n$ respectively. Then*
(4.2)   (i) $p(\lambda)$ *is a left annihilating polynomial for $L[p(\lambda)]$, and a right annihilating polynomial for $L^T[p(\lambda)]$, that is, $p_L(L) = Ip_0 + Lp_1 + L^2p_2 + \cdots + L^n \equiv 0$, and $p_R(L^T) = p_0I + p_1L^T + \cdots + (L^T)^n \equiv 0$.*

(ii) $\{p_L(H[p^m(\lambda)])\}^r = J_m^r \otimes I_n$, $r = 0, 1, \cdots$ *where* $J_m = \begin{bmatrix} 0 & 0 \\ I_{m-1} & 0 \end{bmatrix}$ *and* $\otimes$ *is the right direct product.*

(iii)   *L is invertible in $R_{n \times n}$ if and only if p_0 is invertible in R. In either case*

$$L^{-1} = \begin{bmatrix} -p_1p_0^{-1} & & \vdots & & \\ -p_2p_0^{-1} & & \vdots & & \\ & & \vdots & I_{n-1} & \\ & & \vdots & & \\ -p_{n-1}p_0 & & \vdots & & \\ - & - & -+ & - & - \\ -p_0^{-1} & & \vdots & 0 & \end{bmatrix} = -(Ip_1p_0^{-1} + Lp_2p_0^{-1} + \cdots + L^np_0^{-1})$$

(iv) $G^{-1}L^kG = (L^T)^k \ k = 0, 1, 2, \cdots$ and $\Sigma_{i=0}^n (p_iG^{-1})L^i \equiv 0$.

(v)   If $p_k(\lambda) = p_k + p_{k+1}\lambda + \cdots + \lambda^{n-k} \ k = 0, 1, \cdots, n$, then

$$[I_{n-1} \oplus L(p_{n-1}(\lambda))][I_{n-2} \oplus L(p_{n-2}(\lambda))] \cdots [I_1 \oplus L(p_1(\lambda))]L(p_0(\lambda))G(p_0(\lambda))$$

$$= -\operatorname{diag}[p_0, p_1, \cdots, p_{n-1}],$$

where $\oplus$ denotes the direct sum of two matrices.

(vi) $L^k = \begin{bmatrix} 0 & A_k \\ I_{n-k} & B_k \end{bmatrix}$, where $A_k = L(q_0) \cdots L(q_{k-1})$ and

$$q_i(\lambda) = p_0 + p_1\lambda + \cdots + p_{k-i-1}\lambda^{k-i-1} + p_{n-i}\lambda^{k-i} + \cdots$$

$$\cdots + p_{n-1}\lambda^{k-1} + \lambda^k, \qquad i = 0, \cdots, k-1.$$

Moreover,

$$B_k = \sum_{i=0}^{k-2} C^i Q_i L(q_{i+1}) \cdots L(q_{k-1}) + C^{k-1}Q_{k-1},$$

with $C = [e_2, e_3, \cdots, e_{n-k}, e_1]$, $Q_i = [0, \cdots, 0, p_i]$ and $p_i^T = -[p_k, \ p_{k+1}, \cdots,$
$p_{n-i-1}, \ p_{k-i}, \cdots, \ p_{k-2}, p_{k-1}] \ i = 0, \cdots, k-1$.

(vii) $\lambda I - L[p(\lambda)]$ is equivalent over $R_{n \times n}(\lambda)$ to $\begin{bmatrix} p(\lambda) & 0 \\ 0 & I \end{bmatrix}$;

$\lambda I - H[p^m(\lambda)]$ is equivalent over $R_{n \times n}(\lambda)$ to $\begin{bmatrix} p^m(\lambda) & 0 \\ 0 & I \end{bmatrix}$.

(viii) $L[p^m(\lambda)]$ is similar over $R_{n \times n}$ to $H[p^m(\lambda)]$.

PROOF.   (i) For a field this is a well-known result stating that a companion matrix is non-derogatory with characteristic and minimal polynomials both equal to $p(\lambda)$. In the case of a non-commutative ring the order of multiplication becomes important. Indeed, it is easily seen that for $n = 2$, $p_0I + p_1L + L^2 \neq 0$, that is, $p(\lambda)$ may not be a right annihilating polynomial for $L(p)$. If we write $L = [e_2, e_3, \cdots, e_n, p] = J + E$, where J is given as in (ii) and, $E = [0, \cdots, 0, p]$ with $p^T = -[p_0, p_1, \cdots, p_{n-1}]$,   then $L^k = [e_{k+1}, e_{k+2}, \cdots, e_n, p, Lp, \cdots, L^{k-1}p] \ k = 0, 1, \cdots, n-1$. It is now easily seen by induction that

(4.3) $L^kp = -(L^{k-1}p)p_{n-1} - (L^{k-2}p)p_{n-2} - \cdots - pp_{n-k} + J^kp$

from which, on equating columns, we obtain

$$\{Ip_0 + Lp_1 + \cdots + L^n\}e_i = 0 \qquad i = 1, \cdots, n.$$

The corresponding result for the upper companion matrix $L^T(p(\lambda))$ follows along similar lines or by transposition of the first result (blockwise transposition in case of $R = \mathcal{F}_{n \times n}$).

(ii)   The proof given in (Hartwig (1974)) goes over to the ring case if one uses left functional values and applies the result of (i). It should be noted that $p^m(\lambda)$ is *not* a left annihilating polynomial for H since $\Gamma_H[p^m(\lambda)] \neq [p_L(H)]^m$.

(iii)   This follows immediately from Corollary 3a and part (i). It is further clear that L^{-1} is equivalent to another companion matrix.

(iv)   It may be verified directly that over any ring with unity

$$(4.4) \qquad GL^T = LG = \begin{bmatrix} -p_0 & & & & 0 & & \\ \hline & & p_2 & p_3 & \cdots & p_n \\ 0 & & p_3 & & & \\ & & \vdots & & p_n & \\ & & \vdots & & & \\ & & p_n & & & 0 \end{bmatrix}$$

from which the results immediately follow.

(v)   From (iv) we see that

$$(4.5) \qquad L[p_0(\lambda)]G[p_0(\lambda)] = [-p_0 \oplus G[p_1(\lambda)]],$$

which, when used recursively, yields the desired diagonalization of $G[p_0(\lambda)]$.

(vi)   If E_i denotes the permutation matrix corresponding to interchanging columns $k-i$ and $n-i$, $i = 0, 1, \cdots, k-1$, then it is easily proven by induction that

$$(4.6) \qquad (E_r^{-1} \cdots E_1^{-1} E_0^{-1} L E_0 E_1 \cdots E_r) E_{r+1} = \left[\begin{array}{c|c} L[q_{r+1}(\lambda)] & 0 \\ \hline Q_{r+1} & C \end{array} \right],$$

$$r = 0, 1, \cdots, k-1,$$

where C is the cyclic matrix $[e_2, \cdots, e_n, e_1]$ and $Q_{r+1} = [0, 0, \cdots, p_{r+1}]$, with $p_{r+1}^T = -[p_k, p_{k-1}, \cdots, p_{n-r-2}, p_{k-r-1}, \cdots, p_{k-2}, p_{k-1}]$. Hence

$$L^k E_0 E_1 \cdots E_{k-1} = (L E_0)(E_0^{-1} L E_0 E_1) \cdots$$

$$\cdots (E_{k-1}^{-1} \cdots E_0^{-1} L E_0 E_1 \cdots E_{k-1}) = \begin{bmatrix} X & 0 \\ Y & C \end{bmatrix},$$

where

$$X = L(q_0) \cdots L(q_{k-1})$$

and

$$Y = \sum_{i=0}^{k-2} C^i Q_i L(q_{i+1}) \cdots L(q_{k-1}) + C^{k-1} Q_{k-1}.$$

But

$$E_0 \cdots E_{k-1} = \begin{bmatrix} 0 & C^k \\ I_k & 0 \end{bmatrix}$$

and so

$$L^k = \begin{bmatrix} X & 0 \\ Y & C^k \end{bmatrix} \begin{bmatrix} 0 & I_k \\ C^{-k} & 0 \end{bmatrix} = \begin{bmatrix} 0 & X \\ I & Y \end{bmatrix}.$$

(vii) The derivation of equations (3.13) and (4.8) in (Hartwig (1975)) also hold for an arbitrary ring with unity.

(viii) This result is (as in the case of a field) a consequence of part (vii) and the two versions of the Frobenius theorem, (MacDuffee (1946), Theorem 29.2–3 and Cullen (1972), Theorem 6.15), generalized to rings with unity. We shall state it for completeness and clarity.

THEOREM 4. *Let R be a ring with unity 1 and let $A(\lambda) = A_0 + A_1\lambda$, $B(\lambda) = B_0 + B_1\lambda$ be two binomials in $R_{n\times n}(\lambda)$. Suppose that either A_1 and B_1 are invertible or that B_1 is invertible and $R_{n\times n}$ is finite, that is $AB = I \Rightarrow BA = I$. Then $A(\lambda)$ is equivalent to $B(\lambda)$ over $R_{n\times n}(\lambda)$ if and only if there exist invertible P_0, Q_0 in $R_{n\times n}$ such that $P_0 A(\lambda)Q_0 = B(\lambda)$.*

PROOF. If $R_{n\times n}$ is finite, such as is the case for commutative or unit regular rings (Henriksen (1973)), and B_1^{-1} exists, then the proof of Theorem 6.15 in Cullen (1972) goes through without modification. Indeed, if $P_0 A_1 Q_0 = B_1$ then $P_0(A_1 Q_0 B_1^{-1}) = I = (B_1^{-1} P_0 A)Q_0$ so that P_0^{-1} and Q_0^{-1} both exist. When R is arbitrary, and A_0^{-1}, B_1^{-1} exist then we may combine the proofs of the two versions to establish, in the notation of Cullen (1972) that

    (i) $P_0 A(\lambda)Q_0 = B(\lambda)$ (ii) $R_0 B(\lambda)S_0 = A(\lambda)$

    (iii) $A(\lambda)Q_0 = R_0 B(\lambda)$ (iv) $I = R_0 P_0 = S_0 Q_0$,

where $P(\lambda) = B(\lambda)P_1(\lambda) + P_0$, $Q(\lambda) = Q_1(\lambda)B(\lambda) + Q_0$, $P^{-1}(\lambda) = A(\lambda)R_1(\lambda) + R_0$, and $Q^{-1}(\lambda) = S_1(\lambda)A(\lambda) + S_0$. It is then clear that $A(\lambda)Q_0 S_0 = A(\lambda)$, $P_0 R_0 B(\lambda) = B(\lambda)$ implying that $R_0 = P_0^{-1}$ and $S_0 = Q_0^{-1}$.

COROLLARY 7. *If R is any ring with unity 1, then A is similar to B over $R_{n\times n}$ if and only if $(\lambda I - A)$ is equivalent to $\lambda I - B$ over $R_{n\times n}(\lambda)$.*

THEOREM 9. *Let R be a ring with unity 1, and let $L \in R_{n\times n}$ be the lower companion matrix of $p(\lambda)$. Then L has a group inverse if and only if there exist solutions x and y to*

    (i) $[p_0, p_1] \begin{bmatrix} x \\ y \end{bmatrix} = -1 = [x, y] \begin{bmatrix} p_0 \\ p_1 \end{bmatrix}$

(4.7)   (ii) $p_0 y = 0$

        (iii) $[x, y] \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} y = 0.$

In either case

$$
L^* = \begin{bmatrix}
x_1 & . & y_1 & \vdots & 0 \cdots \cdots 0 \\
\vdots & & \vdots & \vdots & I_{n-2} \\
x_{n-1} & & y_{n-1} & \vdots & \\
\text{---}\text{---}\text{---}\text{---}\text{---}\text{---} & & & \text{---}\text{---}\text{---}\text{---} \\
x & & y & \vdots & 0 \cdots \cdots 0
\end{bmatrix},
$$

where $y_i = \delta_{i1} + p_i y$ *and* $x_i = y_{i+1} + p_i x$, $i = 1, \cdots, n-1$.

PROOF. Suppose first that L^* exists and has columns $[x_1, \cdots, x_n]$. The commutivity of L and L^* implies first of all that

$$LL^* = [x_2, x_3, \cdots, x_n, LL^* e_n].$$

If we set $p = -[p_0, \cdots, p_{n-1}]^T$, then from $LL^* L = L$ we obtain

(4.8) $L^* = [x_1, x_2, e_2, \cdots, e_{n-1}], \quad LL^* e_n = e_n, \quad \text{and} \quad LL^* p = p.$

Using this form of L^* again in the commutivity relation we arrive at the Consistency Conditions

(4.9)
$$
\begin{aligned}
&\text{(a)} \quad Lx_2 = e_2 \\
&\text{(b)} \quad Lx_1 = x_2 \\
&\text{(c)} \quad L^* p = e_n.
\end{aligned}
$$

The last identity $L^* LL^* = L^*$ shows in addition that

(4.9)   (d) $L^* x_2 = x_1.$

On writing $x_1 = [x_1, \cdots, x_n]^T$ and $y_1 = [y_1, \cdots, y_n]^T = x_2^T$ one sees from (5.9a) that

(4.10) $p_0 y_n = 0 \quad \text{and} \quad y_i = \delta_{i1} + p_i y_n \quad i = 1, \cdots, n-1,$

while from (5.9b) one gets

(4.11) $p_0 x_n + p_1 y_n = 1 \quad \text{and} \quad x_i = y_{i+1} + p_i x_n \quad i = 1, \cdots, n-1.$

The Consistency Conditions (5.9c) reduce to

(4.12)
$$
\begin{aligned}
&\text{(a)} \quad x_1 p_0 + y_1 p_1 = 0 \\
&\text{(b)} \quad x_i p_0 + y_i p_i = -p_i \qquad i = 2, \cdots, n-2 \\
\text{and} \quad &\text{(c)} \quad x_n p_0 + y_n p_1 = -1,
\end{aligned}
$$

while (5.9d) yields on using (5.10)

$$\text{(a)} \quad x_1 p_1 y_n + y_1 p_2 y_n = 0$$

(4.13) $\qquad \text{(b)} \quad x_i p_1 y_n + y_i p_2 y_n = - y_{i+1} \qquad i = 2, \cdots, n-1$

$$\text{and} \quad \text{(c)} \quad x_n p_1 y_n + y_n p_2 y_n = 0.$$

The conditions of (5.7) are thus clearly satisfied by x_n and y_n. Conversely suppose that x_n and y_n satisfy (5.7) and suppose that we define y_i and x_i according to (5.10–11). Then all that remains is to verify that the Consistency Conditions (5.12a–b) and (5.13a–b) are also satisfied, since then by the uniqueness of $M^\#$ these must yield the group inverse of M. This verification is lengthy but straightforward and will be omitted. It is crucial, however, to use the following identities which are obtained from (5.7) on pre and postmultiplication by p_0 and p_1 respectively:

$$\text{(a)} \quad p_0 = p_0(-x_n)p_0 \qquad \text{(b)} \quad y_n = y_n(-p_1)y_n$$

(4.14) $\qquad \text{(c)} \quad p_1 y_n p_0 = 0 \qquad\qquad \text{(d)} \quad y_n p_0 = 0$

$$\text{and} \quad \text{(e)} \quad p_0 x_n p_1 = p_1 x_n p_0 = -p_1 - p_1 y_n p_1.$$

Hence when $L^\#$ exists it is not only necessary that p_0 and y_n be regular elements (in the sense of Von Neumann) in R, but also that there exists a 1-inverse of p_0 which obeys

(4.15) $$p_0 p_0^- p_1 = p_1 p_0^- p_0.$$

We remark further that if p_0 ($\neq 0$) is *not* a divisor of zero then $L^\#$ exists if and only if p_0^{-1} and hence L^{-1} exists, while if $p_0 = 0$ [$p_1 = 0$] then $L^\#$ exists if and only if p_1^{-1} [p_0^{-1}] exists.

In the remaining part of this paper let R be a skew-field (division ring) $\mathscr{D}$. We may now use the concept of *rank*, which equals the common row and column rank of a matrix (Jacobson (1953), p. 46), to establish necessary and sufficient conditions for $L^\#$ to exist. Annihilating polynomials, however, are of little use unless $\mathscr{D}$ is commutative.

We first observe that for a square matrix over a skew field

(4.16) $$A^\# \text{ exists if and only if rank } A = \text{rank } A^2.$$

This follows exactly as in the field case from Sylvester's inequality (Jacobson (1953), p. 46), generalized to $\mathscr{D}$, or from Fitting's core-nilpotent decomposition which, as well as the Frobenius Canonical form, also go over to $\mathscr{D}_{n\times n}$ (Jacobson (1943), pp. 50–51; Jacobson (1953), p. 47). In particular, the Drazin inverse A^d *always* exists over $\mathscr{D}$. The conditions of Proposition 9 can be weakened considerably for the case where L is a *block* companion matrix over $\mathscr{D}$, partitioned as

$$L = \begin{bmatrix} 0 & -P_0 \\ I_{(n-1)m} & Q \end{bmatrix} \quad \text{where} \quad Q = -\begin{bmatrix} P_1 \\ \vdots \\ P_{n-1} \end{bmatrix} \quad \text{and} \quad P_i \in \mathcal{D}_{m \times m}.$$

Now rank L = rank I + rank P_0, while by Proposition 7 (vi)

$$\text{rank } L^2 = \text{rank } I_{(n-2)m} + \text{rank} \begin{bmatrix} 0 & -P_0 \\ I_m & -P_1 \end{bmatrix} \begin{bmatrix} 0 & -P_0 \\ I_m & -P_{n-1} \end{bmatrix}$$

$$= (n-2)m + \text{rank} \begin{bmatrix} P_0 & 0 \\ P_1 & P_0 \end{bmatrix}.$$

Thus

$$\text{rank } L = \text{rank } L^2 \Leftrightarrow \text{rank} \begin{bmatrix} P_0 & 0 \\ P_1 & P_0 \end{bmatrix} = \text{rank} \begin{bmatrix} P_0 & 0 \\ 0 & I_m \end{bmatrix}.$$

This may either be calculated by using the rank formula for partitioned matrices (Hartwig (1976), Equation 4.19), which generalizes also to $\mathcal{D}$, or by noting that

$$\begin{bmatrix} P_0 & 0 \\ P_1 & P_0 \end{bmatrix} = \begin{bmatrix} P_0 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ P_1 & P_0 \end{bmatrix}.$$

The latter shows that $L^\#$ exists if and only if $\begin{bmatrix} P_0 & 0 \\ P_1 & P_0 \end{bmatrix}$ and $\begin{bmatrix} P_0 & 0 \\ 0 & I \end{bmatrix}$ have the same range, that is,

$$\begin{bmatrix} P_0 & 0 \\ P_1 & P_0 \end{bmatrix} \begin{bmatrix} X_1 & X_3 \\ X_2 & X_4 \end{bmatrix} = \begin{bmatrix} P_0 & 0 \\ 0 & I \end{bmatrix}$$

for some X_i. This implies that $P_0 X_3 = 0$ and $P_1 X_3 + P_0 X_4 = I_m$. Conversely, if such solutions X_3, X_4 exist, then

$$\begin{bmatrix} P_0 & 0 \\ P_1 & P_0 \end{bmatrix} \begin{bmatrix} I - X_3 P_1 & X_3 \\ -X_4 P_1 & X_4 \end{bmatrix} = \begin{bmatrix} P_0 & 0 \\ 0 & I_m \end{bmatrix},$$

which shows that the first and last matrices have the same range and hence the same rank. A similar result follows on using $\begin{bmatrix} 0 & P_0 \\ I & 0 \end{bmatrix}$. As remarked earlier the companion matrix $L[p(\lambda)]$ has a group inverse over $\mathcal{D}$ if and only if either $p_0 \neq 0$ or $p_0 = 0$ and $p_1 \neq 0$. Now suppose that $p(\lambda) = p_k \lambda^k + \cdots + \lambda^n$, $p_k \neq 0$, $k \geq 0$. Then by (4.2i) the matrix equation $L^{k+1} X = L^k$ has a left polynomial solution which implies, on using rank, that $Y L^{k+1} = L^k$ also has a solution (Jacobson (1943), Theorem 9, p. 51), and hence $i(L) = k$. In general $L^d = L^k X^{k+1}$ will not be a left polynomial in L, however. We may now partition L

5

$$L = \begin{bmatrix} J & 0 \\ N & W \end{bmatrix} \begin{matrix} \}k \\ \}n-k \end{matrix},$$

where $J = [\delta_{i+1,i}]$, $N = E_{1k}$ and $W = L[p_k(\lambda)]$ with $p_k(\lambda)$ as given in (4.2v). Since J is nilpotent and W is invertible we may use (2.0) to calculate the Drazin inverse of L. Indeed, this shows independently that $i(L) = k$ and that

$$(4.17) \qquad L^d = \begin{bmatrix} 0 & 0 \\ X & W_k^{-1} \end{bmatrix}, \qquad LL^d = \begin{bmatrix} 0 & 0 \\ W_k X & I \end{bmatrix},$$

where

$$W_k^{-1} = \left[\begin{array}{c|c} \begin{matrix} -p_{k+1}p_k^{-1} \\ \cdot \\ \cdot \\ \cdot \\ -p_{n-1}p_k^{-1} \end{matrix} & I_{n-k-1} \\ \hline -p_k^{-1} & 0 \end{array} \right] \quad (n-k)\times(n-k)$$

and $X = W_k^{-k-1}Z_k$ with $Z_k = W_k^{k-1}N + W_k^{k-2}NJ + \cdots + NJ^{k-1}$. Whence

$$X = W_k^{-2}N + W_k^{-3}NJ + \cdots + W_k^{-k-1}NJ^{k-1} \in \mathcal{D}_{(n-k)\times k},$$

$$(4.18) \qquad = [W_k^{-k-1}e_1, W_k^{-k}e_1, \cdots, W_k^{-3}e_1, W_k^{-2}e_1]$$

$$= [W_k^{-k}p, W_k^{-k+1}p, \cdots, W_k^{-1}p]$$

where

$$p = -[p_{k+1}p_k^{-1}, \cdots, p_{n-1}p_k^{-1}, p_k^{-1}]^T.$$

In particular, when $k = 1$,

$$(4.19) \qquad L^* = \begin{bmatrix} 0 & 0 \\ W_1^{-2}e_1 & W_1^{-1} \end{bmatrix} \qquad LL^* = \begin{bmatrix} 0 & 0 \\ W_1^{-1}e_1 & I_{n-1} \end{bmatrix}.$$

Now let us return to the hypercompanion matrix $H[p^m(\lambda)]$, where $i(L) = k$. We first observe that while (2.18) implies that $k \leq i(H) \leq mk$, the similarity of H and $L(p^m)$ over $\mathcal{D}$ shows that in fact $i(H) = mk$. Next, we see that by Theorem 1 H^* exists if and only if L^* exists and $(I - LL^*)(NL^*)^j N(I - LL^*) = 0$ $j = 0, 1, \cdots, m - 1$. Substituting from (4.19) we see that

$$(I - LL^*)N(I - LL^*) = \left[\begin{array}{c|c} \begin{matrix} p_1^{-1} \\ p_2p_1^{-1} \\ \cdot \\ \cdot \\ p_1^{-1} \end{matrix} & 0 \end{array} \right] \neq 0,$$

and hence we may conclude that over an arbitrary skew field

(4.20) H^* exists $\Leftrightarrow$ L^{-1} exists $\Leftrightarrow$ $p_0 \neq 0$.

This may also be seen from the similarity of H and $L(p^m)$. Indeed, since $p^m = \lambda^k(p_k^m + \cdots)$, $\{L(p^m)\}^*$ exists if and only if $mk = 0$ or 1, that is if and only if $k = i(L) = 0$. This generalizes the corresponding result for Jordan matrices as remarked at the end of Section 3. Since L^d always exists over $\mathcal{D}$, we may apply Proposition 6 to the hypercompanion matrix $H[p^m(\lambda)]$ to give

(4.21) $H^d = I \otimes B_0 + J \otimes B_1 + \cdots + J^{n-1} \otimes B_n,$

where $B_0 = L^d$,   and $B_1 = -L^d N L^d + (I - LL^d)N(L^d)^{mk+1} + (L^d)^{mk+1}N(I - LL^d)$, in which we substitute from (4.19). This in turn will be similar to $L(p^m)^d$ as given by (4.17).

Added in proof. After this manuscript had been prepared, the authors learned that a formula similar to (2.0) had been developed by Meyer and Rose (submitted) for matrices over a field.

References

A. Ben-Israel and T. N. Greville (1974), *Generalised Inverses Theory and Applications* (Wiley, New York).

Randall E. Cline (1965), 'An application of representations for the generalised inverse of a matrix' (*MRC Technical Report*, **592**).

Charles G. Cullen (1972), *Matrices and Linear Transformations*. Second Edition (Addison–Wesley, Reading, Massachusetts; Menlo Park, California; London; Don Mills, Ontario).

M. P. Drazin (1958), 'Pseudo-inverses in associated rings and semigroups', *Amer. Math. Monthly* **65**, 506–514.

Gertrude Ehrlich (1968), 'Unit-regular rings', *Portugal Math.* **27**, 209–212.

F. R. Gantmacher (1960), *The Theory of Matrices, Volume One* (translated by K. A. Hirsch; Chelsea, New York).

R. E. Hartwig (1974), 'Some properties of hypercompanion matrices', *Industrial Math.* **24**, Part 2, 77–84.

R. E. Hartwig (1975), '$AX - XB = C$, resultants and generalized inverses', *SIAM J. Appl. Math.* **28**, 154–183.

R. E. Hartwig (1976), 'Block generalised inverses', *Arch. Rational Mech. Anal.* **61**, # 3, 197–251.

R. E. Hartwig (to appear), 'Generalised inverses, *EP* elements and associates', *Rev. Roumaine Math. Pures Appl.*

Melvin Henriksen (1973), 'On a class of regular rings that are elementary divisor rings', *Arch. Math. (Basel)* **24**, 133–141.

Nancy M. Huang and Randall E. Cline (1972), 'Inversion of persymmetric matrices having Toeplitz inverses', *J. Assoc. Comput. Mach.* **19**, 437–444.

Nathan Jacobson (1943), *The Theory of Rings* (Mathematical Surveys, II. Amer. Math. Soc., New York City).

Nathan Jacobson (1953), *Lectures in Abstract Algebra. Volume II: Linear Algebra* (Van Nostrand, Princeton, New Jersey; Toronto; New York; London).

Peter Lancaster (1969), *Theory of Matrices* (Academic Press, New York, London).

C. C. MacDuffee (1946), *The Theory of Matrices* (Chelsea, New York).

C. D. Meyer (1975), 'The role of the group generalized inverse in the theory of finite Markov chains', *SIAM. Rev.* **17**, 443–464.

C. D. Meyer and N. J. Rose (submitted), 'The index and the Drazin inverse of block triangular matrices'.

William E. Roth (1952), 'The equations $AX - YB = C$ and $AX - XB = C$ in matrices', *Proc. Amer. Math. Soc.* **3**, 392–396.

Department of Mathematics,
North Carolina State University,
Raleigh, North Carolina 27607,
USA.