Some additive results on Drazin Inverses *

P. Patrício[a]† and R.E. Hartwig[b]

doi: 10.1016/j.amc.2009.05.021

[a]Departamento de Matemática, Universidade do Minho, 4710-057 Braga, Portugal. e-mail: pedro@math.uminho.pt

[b]Mathematics Department, N.C.S.U., Raleigh, NC 27695-8205, U.S.A. e-mail: hartwig@unity.ncsu.edu

Abstract

In this paper, some additive results on Drazin inverse of a sum of Drazin invertible elements are derived. Some converse results are also presented.

Keywords: Drazin inverse, block matrices

AMS classification: 15A09

*Research supported by CMAT - Centro de Matemática da Universidade do Minho, Portugal, by the Portuguese Foundation for Science and Technology – FCT through the research program POCTI, and by Fundação Luso Americana para o Desenvolvimento (project # 273/2008).

†Corresponding author

1 Background

Our aim is to investigate the existence of the Drazin inverse $(p+q)^d$ of the sum $p+q$, where p and q are either ring elements or matrices. The Drazin-inverse is the unique solution to the equations

$$a^{k+1}x = a^k, \; xax = x, \quad ax = xa,$$

for some $k \geq 0$, if any. The minimal such k is called the *index* in(a) of a. If the Drazin inverse exists we shall call the element D-invertible.

An element a is called *regular* if $axa = a$ for some x, and we denote the set of all such solutions by $a\{1\}$.

A ring with 1 is von Neumann (Dedekind) *finite* if $ab = 1 \Rightarrow ba = 1$. Two elements x and y are left(right) orthogonal (LO/RO), if $xy = 0$ (resp. $yx = 0$).

If a is D-invertible, then $a = (a^2a^d) + a(1 - aa^d) = c_a + n_a$ is referred to as the *core-nilpotent* (C-N) decomposition of a.

A knowledge of the D-inverses of p and q *may not* give any information about the existence of the D-inverse of the sum $p+q$, as seen from the case where p and q are both nilpotent. Indeed, if $p = q = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ then $p+q$ is still nilpotent, while if $q = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ then $p+q$ is invertible.

There are two main methods at our disposal, namely we can try to compute $(p+q)^n$ in a *compact* form, or we can use *splittings*.

The former case is based on the fact that the existence of non-negative intergers r and s such that $a^{r+1}x = a^r$ and $a^s = ya^{s+1}$ is equivalent to a is D-invertible. The smallest values of r and s are called the left and right index of a, respectively (see [7]). As shown by Drazin [2], if r and s are finite then $r = s = in(a)$. Furthermore, $a^d = a^m x^{m+1}$ following the proof of the Lemma in [11, page 109], where $m = in(a)$. Indeed, setting $a^d = a^m x^{m+1}$, one can show (i) $aa^d = a^d a$, (ii) $a^{m+1}a^d = a^m$ and (iii) $a^d a a^d = a^d$. We will make use of equalities $y^{m+1}a^{2m+1} = a^m = a^{2m+1}x^{m+1}$.
(i) $aa^d = aa^m x^{m+1} = a^{m+1}x^{m+1} = y^{m+1}a^{m+1} = y^{m+1}a^{2m+1}x^{m+1}a = a^m x^{m+1} a = a^d a$.
(ii) $a^{m+1}a^d = a^{m+1}a^m x^{m+1} = a^m a^{m+1} x^{m+1} = a^{2m+1}x^{m+1} = a^m$.
(iii) Recall that $aa^d = a^d a$ means $a^{m+1}x^{m+1} = a^m x^{m+1} a$, which in turn implies $a^{m+1}x^{m+1}a^m = aa^m x^{m+1} a a^{m-1} = a^{m+2}x^{m+1}a^{m-1} = \cdots = a^{2m+1}x^{m+1}$. Hence, $a^d a a^d = aa^d a^d = a^{m+1}x^{m+1}a^m x^{m+1} = a^{2m+1}x^{m+1}x^{m+1} = a^m x^{m+1} = a^d$.

On the other hand, the key results in the latter direction is given in [8], and states that if p and q have D-inverses, and $pq = 0$, then $(qp)^d$ and $(p+q)^d$ exist and the latter is given by

$$(p+q)^d = (1 - qq^d)[\sum_{r=0}^{k-1} q^r (p^d)^r] p^d + (q^d)[\sum_{r=0}^{k-1} (q^d)^r p^r](1 - pp^d), \tag{1}$$

where $max\{in(p), in(q)\} \leq k \leq \{in(p) + in(q)\}$. Moreover

$$(p+q)(p+q)^d = (1 - qq^d)[\sum_{r=0}^{k-1} q^r (p^d)^r] pp^d + (qq^d)[\sum_{r=0}^{k-1} (q^d)^r p^r](1 - pp^d) + qq^d pp^d \tag{2}$$

This former result is equivalent to the block triangular D-inversion [7]

$$\begin{bmatrix} A & 0 \\ B & D \end{bmatrix}^d = \begin{bmatrix} A^d & 0 \\ X & D^d \end{bmatrix}, \tag{3}$$

where, for $k \geq \{in(A), in(D)\}$,

$$\begin{aligned} X &= -D^d BA^d + (I - DD^d)Y_k(A^d)^{k+1} + (D^d)^{k+1}Y_k(I - AA^d) \\ &= -D^d BA^d + (I - DD^d)R_k(A^d)^2 + (D^d)^2 S_k(I - AA^d), \end{aligned}$$

in which $Y_k = D^{k-1}B + D^{k-2}BA + \cdots + BA^{k-1}$, $R_k = \sum_{t=0} D^t B(A^d)^t$ and $S_k = \sum_{t=0} D^{d^t} BA^t$.

A special application of this gives the interesting result:

Corollary 1.1. If $e^2 = e, f^2 = f$ and $efe = 0 = fef$, then ef, fe and $e + f$ are D-invertible,

$$(e + f)^n = e + f + (n - 1)(ef + fe)$$

and

$$(e + f)^d = e + f - 2(e + fe).$$

Needless to say, this case can be done using *either* powering or by splitting.

Let us end this introductory section by emphasizing a well known result, known as Cline's formula [1] (cf. [7, page 16]), that relates $(ab)^d$ and $(ba)^d$, namely by $(ab)^D = a\left((ba)^D\right)^2 b$.

2 D-inverses via powering

As a first example where powering can be used, we present the case where $a^2 = 0 = b^2$. We have

Proposition 2.1. Suppose a, b and ab are D-invertible and that $a^2 = 0 = b^2$. Then

1. $a + b$ is D-invertible.

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

Proof. Using induction it is easily seen that

$$(a + b)^{2k} = (ab)^k + (ba)^k$$

and

$$(a + b)^{2k+1} = (ab)^k a + (ba)^k b.$$

It is now straight forward to check that $x = a(ba)^d + b(ab)^d$ satisfies the necessary equations $(a+b)^{2k+1}x = (a + b)^{2k}$, $x(a + b)x = x$ and $(a + b)x = x(a + b)$. □

We note in passing that this result takes care of the example of two nilpotent matrices of $p = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $q = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ of index two. This result is not covered by the assumptions that $a^d b = 0 = ab^d = (1 - bb^d)ab(1 - aa^d)$ or that $ab^d = 0 = (1 - bb^d)ab = 0$ of [4] or [3].

When $a^3 = 0 = b^2$, neither the powering method nor the splitting approach seems to yield a tractable path for computing $(a + b)^d$. The example where $a = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ and $b = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ shows that $a + b$ may again be invertible. Note that $ab \neq 0 \neq ba$, so that the basic splitting does not apply. The trouble with the conditions $a^3 = 0 = b^2$ is that the cubic power gives the chain too much freedom, i.e. the expressions for $(a + b)^k$ fail to become periodic. The above example suggests that we *must add* an extra condition to be able to control the number of terms in the powers of the sum. Indeed, we may state

Proposition 2.2. Suppose a, ba^2b are D-invertible and that $a^3 = 0 = b^2 = bab$. Then

1. $a + b$ is D-invertible.
2. $(a + b)^d = (a + b)^m \left[(a^2b)^d a^2 + b(a^2b)^d + ab[(a^2b)^d]\right]^m$ for sufficiently large m.

Proof. If we set $x = a + b$ then it follows by induction for $k = 1, 2, \ldots$ that
(i) $x^{3k} = (a^2b)^k + ab(a^2b)^{k-1}a + (ba^2)^k$;
(ii) $x^{3k+1} = (a^2b)^k a + ab(a^2b)^{k-1}a^2 + b(a^2b)^k$;
(iii) $x^{3k+2} = a^2(ba^2)^k + ab(a^2b)^k + b(a^2b)^k a$.

This shows a 3 term periodicity.

We now may verify directly that

$$(a + b)^{3k+1}u = (a + b)^{3k}$$

and

$$v(a + b)^{3k+1} = (a + b)^{3k-2}$$

where

$$u = (a^2b)^d a^2 + b(a^2b)^d a + ab(a^2b)^d$$

and

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

These ensure that $a + b$ is D-invertible and is given by $(a + b)^d = (a + b)^{3k}u^{3k}$ for sufficiently large k. □

In the next section we shall use a suitable splitting to improve on this result.

3 Splittings

As always our starting point for the splitting approach is the factorization $a+b = \begin{bmatrix} 1 & b \end{bmatrix} \begin{bmatrix} a \\ 1 \end{bmatrix}$. Using Cline's formula [1], we may write

$$(a+b)^d = \begin{bmatrix} 1 & b \end{bmatrix} (M^d)^2 \begin{bmatrix} a \\ 1 \end{bmatrix}, \tag{4}$$

where

$$M = \begin{bmatrix} a \\ 1 \end{bmatrix} \begin{bmatrix} 1 & b \end{bmatrix} = \begin{bmatrix} a & ab \\ 1 & b \end{bmatrix}, \ M^2 = \begin{bmatrix} a^2+ab & a^2b+ab^2 \\ a+b & ab+b^2 \end{bmatrix} \tag{5}$$

and

$$M^3 = \begin{bmatrix} a^3+a^2b+aba+ab^2 & a^3b+a^2b^2+abab+ab^3 \\ a^2+ab+ba+b^2 & a^2b+ab^2+bab+b^3 \end{bmatrix}. \tag{6}$$

There are two approaches that we can take, namely we can compute M^d and then square the result, or we can directly compute $(M^2)^d$ or $(M^3)^d$. We shall start by using the second approach.

Our first result is

Theorem 3.1. Supose that a^2+ab and $ab+b^2$ are D-invertible, and that $a^2b+ab^2=0$. Then $a+b$ is D-invertible with

$$(a+b)^d = \left(a^2+ab\right)^d a + b\left(ab+b^2\right)^d + bXa \tag{7}$$

where

$$\begin{aligned} X &= -\left(ab+b^2\right)^d (a+b)\left(a^2+ab\right)^d + \left[1-\left(ab+b^2\right)\left(ab+b^2\right)^d\right] Y_k \left[\left(a^2+ab\right)^d\right]^{k+1} + \\ &\quad + \left[\left(ab+b^2\right)^d\right]^{k+1} Y_k \left[1-\left(a^2+ab\right)\left(a^2+ab\right)^d\right], \\ Y_k &= \sum_{r=0}^{k-1} \left(ab+b^2\right)^{k-r-1} (a+b)\left(ab+a^2\right)^r \end{aligned}$$

and $in\{a^2+ab\}, in\{b^2+ab\} \le k \le in\{a^2+ab\} + in\{a^2+ab\}$.

Proof. The matrix M^2 collapses to $M^2 = \begin{bmatrix} A & 0 \\ B & D \end{bmatrix}$ where $A = a^2+ab$, $B = a+b$ and $D = ab+b^2$. We may now use equations (3) and (4) to compute the desired D-inverse as $(a+b)^d = \begin{bmatrix} 1 & b \end{bmatrix} \begin{bmatrix} A^d & 0 \\ X & D^d \end{bmatrix} \begin{bmatrix} a \\ 1 \end{bmatrix} = A^d a + bD^d + bXa$. □

Let us now turn to some of the simplifications.

Corollary 3.1. Suppose that a, b, ab, a^2+ab and $ab+b^2$ are D-invertible, and that $a^2b = 0 = ab^2$. Then $a+b$ has a D-inverse as given in (7) which can be expressed in terms of a^d, b^d and $(ab)^d$.

Proof. Since $a^2(ab) = 0$ and $ab(b^2) = 0$, we may use equation (3) to compute the D-inverses, in terms of $a^d, b^d, (ab)^d$ and $(ba)^d$.

First we have,

$$\left(a^2 + ab\right)^k = \sum_{r=0}^{k} (ab)^r \left(a^2\right)^{k-r},$$

$$\left(ab + b^2\right)^k = \sum_{r=0}^{k} \left(b^2\right)^{k-r} (ab)^r,$$

$$\left(a^2 + ab\right)^k b = 0 = a\left(ab + b^2\right)^k, \text{ for } k = 1, 2, ...,$$

and thus $a\left(a^2 + ab\right)^d = 0$.

Using left orthogonality we have in addition, for $A = a^2 + ab$,

$$A^d = \left(a^2 + ab\right)^d = \left[1 - (ab)(ab)^d\right] U_1 \left(a^2\right)^d + (ab)^d U_2 \left(1 - aa^d\right)$$

and

$$AA^d = \left[1 - (ab)(ab)^d\right] U_1 \left(aa^d\right) + (ab)(ab)^d U_2 \left(1 - aa^d\right) + (ab)(ab)^d aa^d,$$

where

$$U_1 = \sum_{r=0}^{N} (ab)^r \left(\left[a^2\right]^d\right)^r$$

and

$$U_2 = \sum_{r=0}^{N} \left[(ab)^d\right]^r \left(a^2\right)^r,$$

for some large enough N.

Likewise,

$$D^d = \left(ab + b^2\right)^d = \left(1 - bb^d\right) V_1 (ab)^d + \left(b^2\right)^d V_2 \left[1 - (ab)(ab)^d\right],$$

$$DD^d = \left(1 - bb^d\right) V_1 (ab)(ab)^d + bb^d V_2 \left[1 - (ab)(ab)^d\right] + bb^d (ab)(ab)^d,$$

where

$$V_1 = \sum_{r=0}^{K} \left(b^2\right)^r \left[(ab)^d\right]^r$$

and

$$V_2 = \sum_{r=0}^{K} \left[\left(b^2\right)^d\right]^r (ab)^r,$$

for some large K.

These can now be used to obtain

$$bD^d (a + b) a =$$
$$b\left[\left(1 - bb^d\right) V_1 (ab)^d + \left(b^2\right)^d V_2 \left[1 - (ab)(ab)^d\right]\right] (a + b) \left[\left[1 - (ab)(ab)^d\right] U_1 \left(a^2\right)^d + (ab)^d U_2 \left(1 - aa^d\right)\right] a$$

as well as

$$b\left(1-DD^{d}\right)R_{k}\left(A^{d}\right)^{2}a=b\left[1-\left(1-bb^{d}\right)V_{1}\left(ab\right)\left(ab\right)^{d}-bb^{d}V_{2}\left[1-\left(ab\right)\left(ab\right)^{d}\right]-bb^{d}ab\left(ab\right)^{d}\right]R_{k}\left(A^{d}\right)^{2}a$$

and the expression

$$\begin{aligned} b\left(D^{d}\right)^{2}S_{k}\left(1-AA^{d}\right)a &= b\left[\left(1-bb^{d}\right)V_{1}\left(ab\right)^{d}+\left(b^{d}\right)^{2}V_{2}\left[1-\left(ab\right)\left(ab\right)^{d}\right]\right]^{2}S_{k}\times \\ &\times \left[1-\left(1-ab\left(ab\right)^{d}\right)U_{1}\left(aa^{d}\right)-ab\left(ab\right)^{d}U_{2}\left(1-aa^{d}\right)-ab\left(ab\right)^{d}aa^{d}\right]a. \end{aligned}$$

These expressions, including R_k and S_k, only use a^d, b^d and $(ab)^d$ via $\left(a^2+ab\right)^d$ and $\left(ab+b^2\right)^d$. □

We next present a useful Lemma.

Lemma 3.1. If $e^2 = e$, $eb = 0$ and b^d exists, then

1. $eb^d = 0$;
2. $(be)^d = 0$;
3. $[b(1-e)]^d = b^d(1-e)$;
4. $b(1-e)[b(1-e)]^d = bb^d$.

Proof. This is left as an exercise. □

It should be noted that a parallel result follows when $af = 0$ with $f^2 = f$.

We now recall the core-nilpotent and Pierce decompositions:

$$a = c_a + n_a \text{ and } b = ebe + eb(1-e) + (1-e)be + (1-e)b(1-e) \tag{8}$$

where $c_a = a^2a^d$ and $n_a = a(1-aa^d)$, if any, and $e^2 = e$.

We may now state

Theorem 3.2. Let a and b be D-invertible with $a^db = 0 = ab^d$. If in addition either $(1-bb^d)ab(1-aa^d) = 0$ or $b(1-bb^d)a(1-aa^d) = n_an_b = 0$, then $a+b$ is D-invertible.

Proof. Let $e = aa^d$ and $f = bb^d$. Then $eb = 0 = af$. We may now split a and b as $a = fa + (1-f)a$ and $b = be + b(1-e) = b_1 + b_2$. By Lemma 3.1, $b_2^d = b^d(1-e)$, $b_2b_2^d = bb^d$, $b_2^2b_2^d = b^2b^d$ and $b_2(1-b_2b_2^d) = b(1-e)(1-f) = b(1-e-f)$.

We now write $a+b = (c_a+b_1)+(n_a+b_2) = x+y$, in which $xy = 0$, on account of $eb = 0 = af$ and Lemma 1.

The D-invertibilty of $a+b$ now follows from equation (1), once we have shown that $x = (c_a+b_1)$ and $y = (n_a+b_2)$ are D-invertible. Since $c_ab_1 = a^2a^d.be = 0$, $c_a^d = a^d$ and $(be)^d = 0$, it is clear from equation (1) that x is D-invertible. On the other hand, to obtain a left orthogonal splitting for y we follow [4] by using a Pierce decomposition for n_a and a CN decomposition for b_2, i.e. let

$$y = (n_a+b_2) = \left[(1-f)n_a + b_2(1-b_2b_2^d)\right] + \left[fn_a + b_2^2b_2^d\right] = u+v.$$

This is a LO splitting because $af = 0 = f(1-f)$. Lastly, to show that u and v are D-invertible, it again suffices to check that we have two LO splittings, and that the summands are D-invertible. In fact, in u we can use

$$(1-f)a(1-e)b(1-e)(1-f) = (1-f)ab(1-e) = 0,$$

or use

$$(1-e)(1-f).(1-e)a = b(1-f)(1-e)a = 0.$$

On the other hand in v we have $fa(1-e)b^2b^d = 0$.

Finally, the four summands fn_a, $(1-f)n_a$, $b(1-e)(1-f)$ and c_b are all D-invertible. In fact the first three summands are nilpotent, while $c_b^d = b^d$. A three fold application of equation (1) gives the actual expression for $(a+b)^d$. □

Remarks

Needless to say, a parallel result holds when $n_b n_a = 0$.

Let us now show that a LO splitting can also be used for our nilpotent example.

Proposition 3.1. Suppose $a^3 = 0 = b^2 = a^2bab = (ab)^3$ and that $(a^2b)^d$ exists. Then $(a+b)^d$ exists and is given by

$$(a+b)^d = ay^d a + bx^d a + by^d a + abx^d + b(ab)^2x^d + (ab)^2[(x^d)^2 + (y^d)^2]a, \tag{9}$$

where $x = a^2b$ and $y = aba$.

Proof. The matrix M^3 of equation (6) reduces to

$$M^3 = \begin{bmatrix} a^2b + aba & abab \\ a^2 + ab + ba & a^2b + bab \end{bmatrix}, \tag{10}$$

which can be split as

$$M^3 = \begin{bmatrix} A & 0 \\ B & D \end{bmatrix} + \begin{bmatrix} 0 & abab \\ 0 & 0 \end{bmatrix} = P + Q, \tag{11}$$

in which $A = a^2b + aba = x + y$ and $D = a^2b + bab = x + n$.

We now note that the assumptions ensure that

$$xn = xab = xb = ax = a^2y = a^2n = abn = bax = xy = 0. \tag{12}$$

We now see that $PQ = 0$, $xy = yx = 0$ and $xn = 0$, so that we have a bi-orthogonal splitting of A and a LO splitting of D. As such both A and D are D-invertible. Consequently,

$$A^k = x^k + y^k, \ A^d = x^d + y^d,$$

and

$$AA^d = xx^d + yy^d.$$

□

It is now convenient here to mention that if x is D-invertible and n is nilpotent of index t with $xn = 0$, then

$$D^d = \left[1 + nx^d + \cdots + n^{t-1}(x^d)^{t-1}\right] x^d \text{ and } DD^d = \left[1 + nx^d + \cdots + n^{t-1}(x^d)^{t-1}\right] xx^d. \tag{13}$$

We shall mainly use the special case where $t = 2$.

Lemma 3.2. Suppose $D = x + n$, where x is D-invertible, $n^2 = 0$ and $xn = 0$. Then

1. $D^k = (x + n)x^{k-1}$, for $k = 1, 2, ...$
2. $D^d = [1 + nx^d]x^d$
3. $DD^d = (x + n)x^d$
4. $(D^d)^k = (1 + nx^d)(x^d)^{k-1} = D^d$, for $k = 2, 3, ...$

The latter shows that D^d is idempotent.

Now, since $PQ = 0 = Q^2$ and P^d exists, we may use equation (1) to obtain

$$(M^3)^d = [I + QP^d]P^d$$

We now can compute the desired D-inverse from

$$(a+b)^d = \begin{bmatrix} 1 & b \end{bmatrix} \begin{bmatrix} a & ab \\ 1 & b \end{bmatrix} (M^3)^d \begin{bmatrix} a \\ 1 \end{bmatrix} = \begin{bmatrix} a+b & ab \end{bmatrix} \left(M^3\right)^d \begin{bmatrix} a \\ 1 \end{bmatrix} \tag{14}$$

Consider $P = \begin{bmatrix} A & 0 \\ B & D \end{bmatrix}$ and $Q = \begin{bmatrix} 0 & abab \\ 0 & 0 \end{bmatrix}$. From equation (1) we know that

$$P^D = \begin{bmatrix} A^d & 0 \\ X & D^d \end{bmatrix}$$

and

$$Q(P^d)^2 = \begin{bmatrix} (ab)^2[XA^d + D^dX] & (ab)^2(D^d)^2 \\ 0 & 0 \end{bmatrix},$$

where $X = -D^dBA^d + R + S$, and

$$R = (1 - DD^d) \left[\sum_{r=0}^{k-1} D^r B (A^d)^r\right] (A^d)^2$$

and

$$S = (D^d)^2 \left[\sum_{r=0}^{k-1} (D^d)^r BA^r\right] (1 - AA^d).$$

Substituting we arrive at

$$(a+b)^d = (a+b)A^d a + (abX)a + (abD^d) + (a+b)(ab)^2 \left[XA^d a + D^d Xa\right] + (a+b)(ab)^2(D^d)^2. \tag{15}$$

Let us now evaluate the six term in this sum using the relations of (12):

1. $D^dBA^d = (1 + nx^d)x^d(a^2 + ab + ba)(x^d + y^d) = 0$ since $xab = 0 = xb = a^2x = a^2y$

2. $(a + b)A^da = (a + b)(x^d + y^d)a = ay^da + bx^da + by^da$

3. $(ab)D^d = ab(1 + nx^d)x^d = abx^d$, as $bn = 0$.

4. $(a + b)(ab)^2(D^d)^2 = (a + b)(ab)^2(1 + nx^d)x^d = b(ab)^2x^d$, as $a(ab)^2 = 0$.

Next we simplify R and S. First we need

Lemma 3.3. If $r \geq 2$, then $D^rB(A^d)^r = 0 = (D^d)^rBA^r$.

Proof. For $r \geq 2$, $D^rB(A^d)^r = (x + n)x^{r-1}(a^2 + ab + ba)[(x^d)^r + (y^d)^r) = 0$, because $a^2x = 0 = a^2y$ and $xab = xba = 0$. Similarly, $(D^d)^rBA^r = (1 + nx^d)(x^d)(a^2 + ab + ba)(x^r + y^r) = 0$. □

We may now simplify R and S.

$$\begin{aligned} R &= (1 - DD^d)(B + DBA^d)(A^d)^2 \\ &= [1 - (x + n)x^d](x + n)(a^2 + ab + ba)[(x^d)^3 + (y^d)^3] \\ &= (ab)[(x^d)^2 + (y^d)^2] + (ba)(y^d)^2 + nab[(x^d)^3 + (y^d)^3]. \end{aligned}$$

Likewise,

$$\begin{aligned} S &= (D^d)^2(B + D^dBA)(I - AA^d) \\ &= (1 + nx^d)x^d(a^2 + ab + ba)[1 - xx^d - yy^d] + (1 + nx^d)(x^d)^2(a^2 + ab + ba)A[1 - xx^d - yy^d] \\ &= (1 + nx^d)x^da^2 \end{aligned}$$

We are now ready for the equalities:

1. $Sa = 0 = abSa = Sx = Sy = Sx^d = Sy^d$

2. $(ab)Ra = ab[(ab)[(x^d)^2 + (y^d)^2] + (ba)(y^d)^2 + nab[(x^d)^3 + (y^d)^3] = (ab)^2[(x^d)^2 + (y^d)^r)$

3. $(a + b)(ab)^2 = b(ab)^2$

4. $(a + b)(ab)^2XA^da = b(ab)^2(R + S)(x^d + y^d)a = b(ab)^2R(x^d + y^d)a = b(ab)^3[(x^d)^2 + (y^d)^2]a = 0$

5. $(a + b)(ab)^2D^dXa = b(ab)^2D^dRa = b(ab)^2(1 + nx^d)x^d\Big[(ab)[(x^d)^2 + (y^d)^2] + (ba)(y^d)^2 + nab[(x^d)^3 + (y^d)^3]\Big]a = 0$

Adding the six terms yields the desired result.

Remarks

1. When $abab = 0$, the last three terms drop out.

2. x^d and y^d are related via $y^d = ab(x^d)^2 a$.

3. For the converse see the next section.

Corollary 3.2. If $a^3 = 0 = b^2 = abab = (ab)^3$ then

$$(a+b)^d = a(aba)^d a + b(a^2b)^d a + b(aba)^d a + ab(a^2b)^d + b(ab)^2(a^2b)^d \tag{16}$$

Let us now return to our previous example, where $a + b = \Omega$.

Example 3.1. Let $a = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ and $b = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$. Then $ab = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $ba = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. This shows that $(ab)^2 = 0 = (ba)^2$. Moreover $y = aba = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = (aba)^d$ and $x = a^2b = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = x^d$. Thus $ay^da = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $bx^da = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $abx^d = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $by^da = 0$. Adding these shows that $(a+b)^d = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \Omega^T$.

4 Converse Results

We shall now assume that $a+b$ is D-invertible, and examine the D-invertibility of the related elements, a, b, ab and ba. We shall present one local result in addition to one global result.

Proposition 4.1. Let $a^3 = 0 = b^2 = a^2bab = baba^2 = 0 = (ab)^3$. If $a+b$ has a Drazin inverse then so do a^2b and aba.

Proof. Using the notation of Proposition 3.1, we see that $nx = 0$. Now if $a+b$ is D-invertible, then the matrices M and M^3 in (5) and (6) are D-invertible, so that $P+Q$ is D-invertible. Now $P = (P+Q)-Q$ is a LO splitting because $PQ = 0 = Q^2$. Consequently, $P^d = \begin{bmatrix} u & w \\ v & z \end{bmatrix}$ exists. This means that for some k,

$$\begin{bmatrix} A^{k+1} & 0 \\ Y_{k+1} & D^{k+1} \end{bmatrix}\begin{bmatrix} u & w \\ v & z \end{bmatrix} = \begin{bmatrix} A^k & 0 \\ Y_k & D^k \end{bmatrix} = \begin{bmatrix} u & w \\ v & z \end{bmatrix}\begin{bmatrix} A^{k+1} & 0 \\ Y_{k+1} & D^{k+1} \end{bmatrix}.$$

This shows that

$$[x^{k+1} + y^{k+1}]u = x^k + y^k$$

and

$$z(x+n)^{k+1} = (x+n)^k \ (k \geq 1).$$

Pre-multiplying the former equation by x then gives $x^{k+2}u = x^{k+1}$, and because $nx = 0$, we also see that the latter reduces to $zx^{k+1} = x^k$. This ensures that x and y are D-invertible. □

We next turn to a global consideration in which we shall assume that our ring is regular and finite.

Proposition 4.2. Given a finite regular ring R and $A = [a_{i,j}]$ a lower triangular matrix over R. If A is group invertible then all $a_{i,i}$ are group invertible.

Proof. Denoting the diagonal element $a_{i,i}$ by a_i, we may write $A = [a_{i,j}] = \begin{bmatrix} a_1 & 0 \\ * & \tilde{A} \end{bmatrix}$. On account of [9] we know that there exists an inner inverse $A^- \in A\{1\}$ such that

$$AA^- = \begin{bmatrix} a_1 a_1^- & 0 \\ * & * \end{bmatrix}.$$

Since $A^\#$ exists,

$$A^2A^- + I - AA^- = \begin{bmatrix} a_1^2 a_1^- + 1 - a_1 a_1^- & 0 \\ * & * \end{bmatrix}$$

is invertible ([12]), from which $a_1^2 a_1^- + 1 - a_1 a_1^-$ is invertible by the finiteness of R. Therefore, $a_1^\#$ exists. Now from [7], we know that the existence of the group inverses for A and a_1, guarantee that $\tilde{A}^\#$ also exists. Repeating this we see that the group invertibility of $\tilde{A}^\#$ implies the group invertibility of a_2. Likewise we obtain the group invertibility of $a_3, \ldots, a_n$. □

Corollary 4.1. Given a finite regular ring R and $A = [a_{i,j}]$ a lower triangular matrix over R. If A is D-invertible then all $a_{i,i}$ are D-invertible.

Proof. If $k = in\,(A)$ then A^k has a group inverse. From Proposition 4.2, the diagonal elements a_i^k of A^k are group invertible as desired. □

Proposition 4.3. If $pq = 0$ and R is finite regular then p^d, q^d exist if and only if $(p+q)^d$ exists.

Proof. If $p+q$ has a D-inverse in ring R, then $\begin{bmatrix} p+q & 0 \\ 0 & 0 \end{bmatrix}$ has a D-inverse in $R_{2\times 2}$. By Cline's formula, if $\begin{bmatrix} p+q & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & q \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p & 0 \\ 1 & 0 \end{bmatrix}$ has a Drazin inverse, so does $\begin{bmatrix} p & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & q \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} p & pq \\ 1 & q \end{bmatrix} = M$. Since $pq = 0$, M reduces to the lower triangular matrix $\begin{bmatrix} p & 0 \\ 1 & q \end{bmatrix}$. From Corollary 4.1, and bearing in mind R is finite, the diagonal elements of M must have Drazin inverses. □

We are now ready for our converse result.

Theorem 4.1. If R is finite regular, $a^2b = 0 = ab^2$ and $(a+b)^d$ exists then a^d, b^d and $(ab)^d$ exist.

Proof. Again, the existence of $(a+b)^d$ implies the Drazin invertibility of $M = \begin{bmatrix} a & ab \\ 1 & b \end{bmatrix}$. Writing $P = \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix}$ and $Q = \begin{bmatrix} 0 & ab \\ 1 & b \end{bmatrix}$, it is clear from $a^2b = 0$ that $M = P + Q$ with $PQ = 0$. This implies, using Cline's formula [1], that $M = \begin{bmatrix} P & 0 \\ I & Q \end{bmatrix}$ is D-invertible. In other words,

$$M = \left[\begin{array}{cc|cc} a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & ab \\ 0 & 1 & 1 & b \end{array}\right]$$

is D-invertible with index, say, k. Hence, M^{2k} has a group inverse, and because $ab^2 = 0$,

$$M^{2k} = \left[\begin{array}{cc|cc} a^{2k} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline * & * & (ab)^{2k} & 0 \\ * & * & * & (b^2+ab)^k \end{array}\right]$$

which is a lower triangular matrix. Using Proposition 4.2, it follows that $(a^{2k})^\#$, $((ab)^{2k})^\#$, $((b^2+ab)^k)^\#$ exist, which imply the D-invertibility of a, ab and of $b^2 + ab$, respectively. Therefore, P^{2k} is group invertible and Q^{2k} is D-invertible, which ensure the D-invertibility of P and Q. In order to complete the proof, we shall show that the existence of Q^d is sufficient for b to be D-invertible. To this effect let us write $Q = \begin{bmatrix} 0 & ab \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & b \end{bmatrix} = K + W$, where $KW = 0$ since $ab^2 = 0$. We claim that the existence of Q^d ensures that K^d and W^d both exist. Indeed, if $(K+W)^d$ exists and $KW = 0$ then, again by Cline's formula, $Z = \begin{bmatrix} K & KW \\ I & W \end{bmatrix}$ is D-invertible. Since K is a counter-diagonal matrix, its even powers are diagonal matrices. In fact, $K^{2n} = \begin{bmatrix} (ab)^n & 0 \\ 0 & 1 \end{bmatrix}$. Since $(ab)^d$ exists with Drazin index, say, r, then $(ab)^l$ are all group invertible for $l \geq r$. In particular $(ab)^{2r}$ has a group inverse, which means $K^{2r} = \begin{bmatrix} (ab)^r & 0 \\ 0 & 1 \end{bmatrix}$ has a group inverse. Therefore, K has a Drazin inverse. Lastly, since K and Z are D-invertible, it again follows from [7], that W^d exists, ensuring that b is D-invertible. □

We conclude with the observation that if a (and hence all powers of a) has a right (left) inverse and is D-invertible, then a must be a unit.

Acknowledgment

The authors wish to thank an anonymous referee for his/her remarks and corrections.

References

[1] R.E. Cline, An application of representation of a matrix, *MRC Technical Report*, 592, 1965.

[2] M. P., Drazin, Pseudo-inverses in associative rings and semigroups, *Amer. Math. Monthly*, 65 (1958), 506–514.

[3] Dragan S. Djordjević, Yimin Wei, Additive results for the generalized Drazin inverse, *J. Aust. Math. Soc.*, 73 (2002), no. 1, 115–125.

[4] N. Castro González, Additive perturbation results for the Drazin inverse, *Linear Algebra Appl.*, 397 (2005), 279–297.

[5] N. Castro González, J.J. Koliha, New additive results for the g-Drazin inverse, *Proc. Roy. Soc. Edinburgh Sect. A*, 134 (2004), no. 6, 1085–1097.

[6] R.E. Hartwig, J. Luh, On finite regular rings, *Pacific J. Math.*, 69 (1977), no. 1, 73–95.

[7] R.E. Hartwig, J. Shoaf, Group Inverses and Drazin inverse of bidiagonal and triangular Toeplitz matrices, *J. Austral. Math. Soc. Ser. A*, 24 (1977), no. 1, 10–34.

[8] R.E. Hartwig, G. Wang and Y. Wei, Some additive results on Drazin inverses, *Linear Algebra Appl.*, 322 (2001), no. 1-3, 207–217.

[9] P. Patrício, R. Puystjens, About the von Neumann regularity of triangular block matrices. *Linear Algebra Appl.*, 332/334 (2001), 485–502.

[10] P. Patrício, *Uma contribuição para a Teoria Algébrica da Invertibilidade Generalizada*, Ph.D. thesis, Universidade do Minho, 2002.

[11] R. Puystjens, M.C. Gouveia, Drazin invertibility for matrices over an arbitrary ring, *Linear Algebra Appl.* 385 (2004), 105–116.

[12] R. Puystjens, R.E. Hartwig, The group inverse of a companion matrix. *Linear and Multilinear Algebra*, 43 (1997), no. 1-3, 137–150.