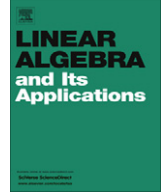




Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa



A class of outer generalized inverses

Michael P. Drazin

Department of Mathematics, Purdue University, West Lafayette, IN 47907-2067, United States

ARTICLE INFO

Article history:

Received 8 October 2010

Accepted 3 September 2011

Available online 13 October 2011

Submitted by E. Zerz

AMS classification:

15A09

06A11

16P99

20M99

Keywords:

Bott–Duffin inverse

Computation of generalized inverses

Exchange ring

Extremal properties

Mitsch partial order

Moore–Penrose generalized inverse

Outer generalized inverses

Potent ring

Pseudo-inverse

Semigroup

Stable range one

Strong π -regularity

Strongly clean ring

Suitable ring

Weighted inverse

W -weighted pseudo-inverse

ABSTRACT

In any $*$ -semigroup or semigroup S , it is shown that the Moore–Penrose inverse $y = a^\dagger$, the author's pseudo-inverse $y = a'$, Chipman's weighted inverse and the Bott–Duffin inverse are all special cases of the more general class of “ (b, c) -inverses” $y \in S$ satisfying $y \in (bSy) \cap (ySc)$, $yab = b$ and $cay = c$. These (b, c) -inverses always satisfy $yay = y$, are always unique when they exist, and exist if and only if $b \in Scab$ and $c \in cabS$, in which case, under the partial order \mathcal{M} of Mitsch, y is also the unique \mathcal{M} -greatest element of the set $X_a = X_{a,b,c} = \{x: x \in S, xax = x \text{ and } x \in (bSx) \cap (xSc)\}$ and the unique \mathcal{M} -least element of $Z_a = Z_{a,b,c} = \{z: z \in S, zaz = z, zab = b \text{ and } caz = c\}$. The above all holds in arbitrary semigroups S , hence in particular in any associative ring R . For any complex $n \times n$ matrices a, b, c , an efficient uniform procedure is given to compute the (b, c) -inverse of a whenever it exists. In the ring case, $a \in R$ is called “weakly invertible” if there exist $b, c \in R$ satisfying $1 - b \in (1 - a)R$, $1 - c \in R(1 - a)$ such that a has a (b, c) -inverse y satisfying $ay = ya$, and it is shown that a is weakly invertible if and only if a is strongly clean in the sense of Nicholson, i.e. $a = u + e$ for some unit u and idempotent e with $eu = ue$.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

In this article we outline a unified theory which encompasses a large class of (and even, in a sense, all) uniquely-defined outer generalized inverses. While we shall be concerned with inverses of several types and at several levels of generality, the ideas arose most directly from two specific and well-known

E-mail address: mdrazin@math.purdue.edu

generalized inverses, the Moore–Penrose inverse a^\dagger and my own pseudo-inverse a' . For the reader's convenience, we first recall their definitions:

Definition 1.1. Given any multiplicative semigroup S , any specified involution of S (i.e. any map $*$: $S \rightarrow S$ satisfying $(ab)^* = b^*a^*$ and $(a^*)^* = a$) and any $a, y \in S$, then (as in [18], or see e.g. [3, pp. 40–49]) y is called a *Moore–Penrose inverse* of a in S if $yay = y$, $aya = a$, $(ay)^* = ay$ and $(ya)^* = ya$.

This y is always unique (whenever it exists), and so is called the Moore–Penrose inverse of a in S ; it is denoted $y = a^\dagger$.

Definition 1.2. Given any semigroup S and any $a, y \in S$, then (as in [7], or see e.g. [3, pp. 163–172]) we call y a *pseudo-inverse* of a in S if $yay = y$, $ay = ya$ and $a^{j+1}y = a^j$ for some $j \in \mathbb{N}$.

This y also is unique when it exists, and so we call it the pseudo-inverse of a in S ; we denote it as $y = a'$.

Thus $y = a^\dagger$ and $y = a'$ both share, by explicit definition, the “outer” inverse property $yay = y$, and moreover the respective arguments used to prove that each of a^\dagger and a' is uniquely determined by a are closely similar; however, in other ways a^\dagger and a' behave very differently.

In Sections 2–4 and 6 below we explain, and in Section 5 exploit, these similarities and differences. We first note two consequences which the respective definitions of each of $y = a^\dagger$ and $y = a'$ both have in common: *Given any suitable associative multiplicative system S , any element $a \in S$, and suitable $b, c \in S$, then y satisfies both*

- (1) $y \in (bSy) \cap (ySc)$ and
- (2) $yab = b$ and $cay = c$.

Explicitly, for $y = a^\dagger$, we take $b = c = a^*$, and have

- (1) $y = yay = (ya)^*y = a^*y^*y \in bSy$, and dually $y \in ySc$, and
- (2) $yab = (ya)^*a^* = (aya)^* = a^* = b$, and dually $cay = c$,

while, for $y = a'$, we take $b = c = a^j$, and have

- (1) $y = yay = y(ay)^j = a^jy^{j+1} \in bSy$, and dually $y \in ySc$, and
- (2) $yab = ya^{j+1} = a^j = b$, and dually $cay = c$.

Note that, for $y = a^\dagger$, there is no need to require that the involution $*$ be “proper” (so that, e.g., for $n \times n$ complex matrices, $*$ can be ordinary transposition, even without complex conjugation). Moreover, by the same two-line argument above and Theorem 2.1 below, the definition and uniqueness of the Moore–Penrose inverse a^\dagger do not at all depend on the property $(a^*)^* = a$, and indeed the map $*$: $S \rightarrow S$ need not even be surjective or injective: all that really matters for (1) and (2) is that $*$: $S \rightarrow S$ should be anti-homomorphic, i.e. should satisfy $(ab)^* = b^*a^*$. The need to assume properness and $(a^*)^* = a$ arises only in applying Theorem 2.2 to prove that a^\dagger exists (e.g. for all $a \in S$ in every strongly π -regular proper $*$ -semigroup S).

The argument for $y = a^\dagger$ extends easily to Chipman’s “weighted inverse” ([5, pp. 114–176], or see e.g. [3, pp. 119–120]), defined by the equations $yay = y$, $aya = a$, $(yav)^* = yav$ and $(way)^* = way$, where v and w can be any given invertible elements of S (for our purposes there is no need to impose any condition corresponding to positive definiteness): it suffices to take $b = (av)^*$ and $c = (wa)^*$. The weighted inverse provides a case where in general $b \neq c$.

Similarly, although with some anomalies, the argument for $y = a'$ extends to Cline and Greville’s [6] “ W -weighted pseudo-inverse” x defined by $xwawx = x$, $awx = xwa$ and $(aw)^{j+1}xw = (aw)^j$ for some j , of which the last two imply that also, dually, $wx(wa)^{j+2} = (wa)^{j+1}$. On taking $b = (wa)^{j+1}$,

$c = (aw)^j$, then $y = wxw$ is the (b, c) -inverse of a . For (2) is immediate, while

$$x = xw(awx) = xw(xwa) = (xw)x(wa) = \cdots = (xw)^j x(wa)^j,$$

so that $y = w(xw)^j x(wa)^j w = y(xw)^j c \in ySc$, and dually $y \in bSy$, which gives (1).

As will be described in Section 3, another choice of b and c (with $b = c$ idempotent) yields the Bott–Duffin inverse. We also introduce (see Definition 3.2) a simpler and more general version of the Bott–Duffin inverse which, at least formally, includes both the Moore–Penrose inverse and the pseudo-inverse.

As conditions on y , the properties (1) and (2) are, as indicated by Example 2.5, very weak, but, even for arbitrary a, b, c , are nevertheless together sufficient to ensure the uniqueness (Theorem 2.1(i)) of y for given a (and given b, c); and Theorem 2.2 provides a very simple necessary and sufficient condition on a, b, c (namely, that $b \in Scab$ and $c \in cabS$) for the existence of such a y . Also, although neither (1) nor (2) mentions the outer property $yay = y$ explicitly, they do together imply it (Theorem 2.1(ii)).

To discuss these matters more formally, we introduce

Definition 1.3. Let S be any semigroup and let $a, b, c, y \in S$. Then we shall call y a (b, c) -inverse of a if both

- (1) $y \in (bSy) \cap (ySc)$ and
- (2) $yab = b, cay = c$.

This terminology reflects the fact that, if a is invertible, then obviously $y = a^{-1}$ satisfies (2), but is mainly motivated by Theorem 2.1 below. For the classical inverse $y = a^{-1}$, defined as usual by $ya = ay = 1$, just take $b = c = 1$.

In Section 4, still for arbitrary semigroups S , by using the partial order \mathcal{M} introduced by Mitsch in 1986 on the set of elements of S , we show in Theorem 4.3 that, for given $a, b, c \in S$, the (b, c) -inverse y of a , when y exists, is always the unique \mathcal{M} -greatest element of the set

$$X_a = X_{a,b,c} = \{x: x \in S, xax = x \text{ and } x \in (bSx) \cap (xSc)\},$$

and also the unique \mathcal{M} -least element of

$$Z_a = Z_{a,b,c} = \{z: z \in S, zaz = z, zab = b \text{ and } caz = c\}.$$

These properties of y open the door to more general alternative versions of the (b, c) -inverse which may be meaningful even when no $y \in S$ satisfies Definition 1.3.

Of course Sections 2–4 all apply, in particular, to arbitrary associative rings, e.g. to the algebra $M_n(F)$ of all $n \times n$ matrices over a field F . While generalized inverses are traditionally, for better or worse, most often studied for $S = M_n(F)$, for the most part nothing more need be said here about this special case. However, as a by-product of the proof of Theorem 2.2, for any $a, b, c \in M_n(\mathbb{C})$, we describe in Remark 2.3 a uniform and simple procedure to compute the (b, c) -inverse y of a which, even for $y = a^\dagger$ and $y = a'$, may be more efficient than other methods currently in use. Also, in connection with Theorem 2.2 (and more generally whenever dealing with conditions of the form $p \in Srp$ and/or $p \in prS$), we note in Remark 2.4 that, given any matrices $a, b, c \in M_n(F)$, then a has a (b, c) -inverse y if and only if

$$\text{rank}(c) = \text{rank}(cab) = \text{rank}(b).$$

Moreover, in Example 2.5 and Remark 2.6, we give a complete description of the (b, c) -inverses y for all $a, b, c \in M_2(F)$ with a singular.

To accommodate non-square matrices, Definition 1.3 extends easily to the case where a is $m \times n$ and b, c, y are $n \times m$ (this corresponds to regarding a and b, c, y as maps in a two-object category rather than as elements of a single semigroup S).

In Section 5 we consider, for an arbitrary associative ring R with 1, elements $y \in R$ satisfying, besides (1) and (2) of Definition 1.3, also (3) $1 - b \in (1 - a)R$, $1 - c \in R(1 - a)$ and (4) $ay = ya$. We call

$a \in R$ weakly invertible whenever there exist $b, c, y \in R$ satisfying (1)–(4), and show in Theorem 5.5 that weak invertibility for $a \in R$ is equivalent to the existence of $y \in R$ such that

$$yay = y, \quad 1 - y \in ((1 - a)R) \cap (R(1 - a)) \text{ and } ay = ya,$$

and also to Nicholson's property of a being “strongly clean”.

In Section 6 we extend Theorem 2.1(i) by assuming that $yay = y$ and replacing (1) and (2) by either of two alternative weaker hypotheses, expressed in terms of right or left annihilator ideals, which guarantee that y is still uniquely determined, and each of which leads to a corresponding generalization of Theorem 4.3.

I am indebted to an anonymous referee for detailed criticisms of my article as originally submitted, which have helped to clarify this revised version, and also for encouraging me to provide matrix examples of the questions discussed.

2. (b, c) -Inverses

Theorem 2.1. *For any semigroup S , and any $a, b, c \in S$,*

- (i) *there can be at most one (b, c) -inverse y of a , and*
- (ii) *any such y must also satisfy $yay = y$.*

Proof. Assume that a, b, c are given and that x and y are two (b, c) -inverses of a . By (1) there are $g, h \in S$ such that $x = xgc$ and $y = bhy$. By (2) we have $b = xab$ and $cay = c$. Then

$$x \stackrel{(1)}{=} xgc \stackrel{(2)}{=} xg(cay) = (xgc)ay \stackrel{(1)}{=} xay,$$

and dually

$$y \stackrel{(1)}{=} bhy \stackrel{(2)}{=} (xab)hy = xa(bhy) \stackrel{(1)}{=} xay.$$

Hence $x = y$, and (ii) also follows. \square

Conversely, as evidence of the weakness of (1) and (2) for unspecified (b, c) , for any given $a, y \in S$, if $yay = y$ then obviously y is itself both the (y, y) -inverse and the (ya, ay) -inverse of a (so that a given $y \in S$ is a (b, c) -inverse of a for some $b, c \in S$ if and only if $yay = y$). Thus, for example, the Moore–Penrose inverse a^\dagger is the (b, c) -inverse of a with three different choices of (b, c) , namely $(b, c) = (a^*, a^*), (a^\dagger, a^\dagger)$ or $(a^\dagger a, aa^\dagger)$, and similarly for a' .

Note that Theorem 2.1 shows that $y = a'$ is independent of j .

Theorem 2.2. *For any given semigroup S , and any given $a, b, c \in S$, there exists at least one (b, c) -inverse y of a if and only if $b \in Scab$ and $c \in cabS$.*

Proof. If y exists, then, by (2) and (1), we have at once $b = yab \in (ySc)ab \subseteq Scab$, and dually also $c \in cabS$.

Conversely, if $b \in Scab$ and $c \in cabS$, i.e. if $b = vcab$ and $c = cabw$ for suitable $v, w \in S$, then

$$b = v(cabw)ab = (vcab)wab = bwab,$$

and dually $c = cawc$, so that $vc = v(cabw) = (vcab)w = bw$.

For $y = vc = bw$, we then have

$$y = (bwab)w = bwa(bw) = bway \in bSy,$$

and dually $y \in ySc$, i.e. this y satisfies (1).

Also $yab = bwab = b$, and dually $cay = c$. Thus y satisfies (2), as required. \square

For example, $a^* = (a^\dagger a^{\dagger*})a^*aa^*$, $a^j = (a')^{j+1}a^jaa^j$, and dually.

Remark 2.3. When S is the F -algebra $M_n(F)$ the argument above not only proves (under appropriate hypotheses) the existence of the (b, c) -inverse y , but also provides an easy way to compute it (when it exists) by purely Gaussian procedures, e.g. as $y = bw$ where w is *any* solution of $c = cabw$. For example, for $y = a'$, while the formal solution $w = (a')^{j+1}$ of $a^j = a^j a a^j w$ is of course not of practical use, it suffices to use *any* solution w of $a^j = a^{2j+1}w$ (where we may replace j by n) and obtain $a' = a^j w$. Likewise, when $F = \mathbb{C}$, although the formal solution $w = a^{\dagger*}a^\dagger$ of $a^* = a^* a a^* w$ is not usable, we can use any solution w of $a^* = a^* a a^* w$ to obtain $a^\dagger = a^* w$ (this being a reasonably efficient and practical way to compute a^\dagger for large n , whereas for a' the need to first compute a^n introduces a factor $(\log n)^2$ into the cost).

Remark 2.4. When $S = M_n(F)$, of course any statement of the form $p \in qS$ (as in Theorem 2.2), for given $p, q \in S$, translates at once as $\text{im}(p) \subseteq \text{im}(q)$ for p, q regarded as maps $p, q: F^n \rightarrow F^n$ (equivalently, the column space of p lies in that of q). Thus, in Theorem 2.2, when $a, b, c \in M_n(F)$, then y exists if and only if $\text{rank}(c) = \text{rank}(cab) = \text{rank}(b)$.

In view of Theorem 2.2, we shall say that the pair (b, c) is *a-compatible* if $b \in Scab$ and $c \in cabS$. For any *a-compatible* pair (b, c) , if b or c is zero then obviously $b = c = y = 0$ (or equivalently $cab = 0$), and for any $a \in S$ this trivial case is always one possibility. For $S = M_2(F)$ and two representative singular $a \in S$ (so that b and c must also be singular), we note next all the possible non-trivial *a-compatible* pairs (b, c) :

Example 2.5. Let $a = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then, to avoid the trivial case $cab = 0$, b must have nonzero second row and c must have nonzero first column, say

$$b = \begin{pmatrix} \lambda\alpha & \lambda\beta \\ \alpha & \beta \end{pmatrix} = \begin{pmatrix} \lambda \\ 1 \end{pmatrix} (\alpha, \beta), \quad c = \begin{pmatrix} \gamma & \gamma\mu \\ \delta & \delta\mu \end{pmatrix} = \begin{pmatrix} \gamma \\ \delta \end{pmatrix} (1, \mu).$$

Then

$$cab = \begin{pmatrix} \gamma \\ \delta \end{pmatrix} (1, \mu) a \begin{pmatrix} \lambda \\ 1 \end{pmatrix} (\alpha, \beta) = \begin{pmatrix} \gamma \\ \delta \end{pmatrix} (\alpha, \beta),$$

and we can choose scalars θ, ϕ, ξ, η satisfying

$$(\theta, \phi) \begin{pmatrix} \gamma \\ \delta \end{pmatrix} = (\alpha, \beta) \begin{pmatrix} \xi \\ \eta \end{pmatrix} = 1,$$

so that

$$b = \begin{pmatrix} \lambda \\ 1 \end{pmatrix} (\alpha, \beta) = \begin{pmatrix} \lambda \\ 1 \end{pmatrix} (\theta, \phi) \begin{pmatrix} \gamma \\ \delta \end{pmatrix} (\alpha, \beta) = \begin{pmatrix} \lambda \\ 1 \end{pmatrix} (\theta, \phi) cab \in Scab$$

and

$$c = \begin{pmatrix} \gamma \\ \delta \end{pmatrix} (1, \mu) = \begin{pmatrix} \gamma \\ \delta \end{pmatrix} (\alpha, \beta) \begin{pmatrix} \xi \\ \eta \end{pmatrix} (1, \mu) = cab \begin{pmatrix} \xi \\ \eta \end{pmatrix} (1, \mu) \in cabS.$$

Thus, for all scalars λ, μ and all $(\alpha, \beta) \neq 0, (\gamma, \delta) \neq 0$, the pair (b, c) as above is *always* a -compatible, and for these a, b, c we also immediately find

$$y = bw = \begin{pmatrix} \lambda \\ 1 \end{pmatrix} (\alpha, \beta) \begin{pmatrix} \xi \\ \eta \end{pmatrix} (1, \mu) = \begin{pmatrix} \lambda \\ 1 \end{pmatrix} (1, \mu) = \begin{pmatrix} \lambda & \lambda\mu \\ 1 & \mu \end{pmatrix}.$$

Similarly, for $a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $b = \begin{pmatrix} \alpha & \beta \\ \lambda\alpha & \lambda\beta \end{pmatrix}$, $c = \begin{pmatrix} \gamma & \gamma\mu \\ \delta & \delta\mu \end{pmatrix}$, we easily verify that (b, c) is always a -compatible, now with $y = \begin{pmatrix} 1 & \mu \\ \lambda & \lambda\mu \end{pmatrix}$.

Thus the restrictions placed by Definition 1.3 on b, c are very weak, in that (for these a) they allow nearly all b, c of rank 1. It is of particular interest (again, for these a) that, subject only to y needing to have 1 as an entry (in a location dictated by a), we can arrange for y to be *any* matrix of rank 1 by choosing b and c appropriately.

Remark 2.6. As an alternative to Example 2.5, one can combine the two cases $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ by considering instead the general 2×2 singular matrix $a = \begin{pmatrix} p \\ q \end{pmatrix} (r, s)$, now with $b = \begin{pmatrix} \lambda \\ \mu \end{pmatrix} (\alpha, \beta)$ and $c = \begin{pmatrix} \gamma \\ \delta \end{pmatrix} (\rho, \sigma)$. Then $cab = \tau \begin{pmatrix} \gamma \\ \delta \end{pmatrix} (\alpha, \beta)$, where the scalar $\tau = (\rho, \sigma) \begin{pmatrix} p \\ q \end{pmatrix} (r, s) \begin{pmatrix} \lambda \\ \mu \end{pmatrix}$. To avoid the trivial case $cab = 0$ we must have $\tau \neq 0$, $\begin{pmatrix} \gamma \\ \delta \end{pmatrix} \neq 0$ and $(\alpha, \beta) \neq 0$, under which conditions (by choosing θ, ϕ, ξ, η much as before) we again of course find that every such pair (b, c) is a -compatible, now with $y = \begin{pmatrix} \lambda \\ \mu \end{pmatrix} (\rho, \sigma)$.

However, the two cases $a = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ in Example 2.5 already cover every singular $a \in M_2(F)$ up to similarity and multiplication by nonzero scalars.

For rings with unity, a -compatibility can be equivalently described in terms of properties of the left and right annihilator ideals $\text{lann}(b)$ and $\text{rann}(c)$:

Proposition 2.7. *Let R be any associative ring with 1. Then, for any given $a, b, c \in R$, the following are equivalent:*

- (i) *the pair (b, c) is a -compatible;*
- (ii) $R = abR \oplus \text{rann}(c) = Rca \oplus \text{lann}(b)$;
- (iii) $R = abR + \text{rann}(c) = Rca + \text{lann}(b)$.

Proof. Obviously (ii) \Rightarrow (iii), and so it suffices to verify that (i) \Rightarrow (ii) and (iii) \Rightarrow (i).

(i) \Rightarrow (ii). If (i) holds then there are $v, w \in S$ such that $b = vcab$ and $c = cabw$. Write $r = 1 - abw$. Then $r \in \text{rann}(c)$, so that, for any $s \in R$, we have $s = (abw + r)s \in abR + \text{rann}(c)$, i.e. $R = abR + \text{rann}(c)$.

Moreover, if $u \in (abR) \cap \text{rann}(c)$, say with $u = abt$, then $bt = (vcab)t = v(cu) = 0$, whence $u = abt = 0$. Thus the sum $abR + \text{rann}(c)$ is direct, and dually also $R = Rca \oplus \text{lann}(b)$.

(iii) \Rightarrow (i). If $R = abR + \text{rann}(c)$ then $c \in cR \subseteq cabR$. Dually, $R = Rca + \text{lann}(b)$ gives $b \in Rcab$. \square

At least when $b = c$, one can generalize known properties of a' by establishing sufficient conditions under which $ay = ya$, and also (cf [7, Theorem 1]) for y to commute with every $d \in S$ such that $ad = da$. However, these results seem somewhat artificial, and to prove them here would be a distraction from our main topic.

3. Bott–Duffin inverses

For $S = M_n(\mathbb{C})$, suitably given $a, e \in S$ with $e^2 = e = e^*$, and regarding e as the orthogonal projection onto the subspace $L = e\mathbb{C}^n$, Bott and Duffin ([4], or see e.g. [3, p. 92, Theorem 16]) in 1953 defined a corresponding element $y \in M_n(\mathbb{C})$ such that

- (i) $y = ey = ye$ and
- (ii) $yae = e = eay$,

which they regarded merely as properties of y rather than as defining y (see below). However, even in any semigroup S , since (i) and (ii) respectively obviously imply our (1) and (2) for $(b, c) = (e, e)$, Theorem 2.1 guarantees that (i) and (ii) can have at most one solution y , while Theorem 2.2 tells us that y exists if and only if $e \in \text{Seae} \cap \text{eaeS}$ (or, for $S = M_n(F)$, equivalently $\text{rank}(eae) = \text{rank}(e)$). Moreover, the restrictions $S = M_n(\mathbb{C})$, $e^* = e$ (and even the presence of any involution $*$ on S) now transpire to have been only irrelevant distractions. The y of (i) and (ii) is called the *Bott–Duffin generalized inverse of a relative to e* , but Bott and Duffin themselves defined it not via (i) and (ii) but instead by the explicit formula $y = e(1 - e + ae)^{-1}$. Since it appears that $1 - e + ae$ might be singular, they regarded y as being well-defined only for choices of a and e such that $1 - e + ae$ happens to be invertible, in which case (as they noted) (i) and (ii) follow easily from their definition of y .

Although (i) and (ii) are expressed in semigroup (i.e. purely multiplicative) language, of course Bott and Duffin's own definition of y uses the ring operations of addition and subtraction. However, as we note next, in rings the invertibility of $1 - e + ae$ is an automatic consequence of (i) and (ii):

Proposition 3.1. *For any associative ring R with 1 and any $a, e, y \in R$ with $e^2 = e$, if (i) and (ii) both hold then $1 - e + ae$ is invertible, with $(1 - e + ae)^{-1} = 1 - ay + y$ and $y = e(1 - e + ae)^{-1}$.*

Dually, also $1 - e + ea$ is invertible, with $(1 - e + ea)^{-1} = 1 - ya + y$, and $y = (1 - e + ea)^{-1}e$.

Proof. By (i) and (ii), it is easy to verify that

$$(1 - e + ae)(1 - ay + y) = 1 = (1 - ay + y)(1 - e + ae),$$

so that the explicit expression $y = e(1 - e + ae)^{-1}$ follows at once from either $y(1 - e + ae) = e$ or $y = e(1 - ay + y)$. \square

Note incidentally that $y(1 - e + ae)y = y$ and also

$$ae(1 - ay + y)ae = ae(ae - a(yae) + yae) = ae(ae - ae + e) = ae,$$

so that y , ae , and dually also ea , are all unit regular.

Bott and Duffin's (i) and (ii) are so close to our (1) and (2) that, by hindsight, the Bott–Duffin inverse must now be recognized as a direct precursor of the (b, c) -inverse, and hence also of the subsequent development of most other known types of uniquely-defined outer generalized inverses.

We next introduce another generalized inverse, intermediate between the Bott–Duffin inverse and the (b, c) -inverse:

Definition 3.2. For any semigroup S and any $a, e, f, y \in S$ with e and f both idempotent, call y a Bott–Duffin (e, f) -inverse of a if

- (i) $y = ey = yf$
- and (ii) $yae = e$ and $fay = f$.

For an example of a case covered by Definition 3.2 but not by the original Bott–Duffin definition, take $e = a^\dagger a$, $f = aa^\dagger$ and $y = a^\dagger$. Note that (as suggested by this example) in Definition 3.2, if y exists then e and f must be *equivalent* idempotents, in the standard sense that there exist $p, q \in S$ with $e = pq$ and $f = qp$: for (i) and (ii) give $e = yae = (yf)ae = y(fae)$, and dually $f = (fae)y$. It is easy to verify from Example 2.5 that nothing like this can hold for b and c in Definition 1.3, since, for each a and any λ, μ , we can arrange that $\text{trace}(b) \neq \text{trace}(c)$.

Clearly a (b, c) -inverse y of a is a Bott–Duffin (b, c) -inverse of a if and only if b and c are both idempotent. By Theorem 2.1(i), y in Definition 3.2 is unique when it exists, and we also immediately have

Proposition 3.3. *For any semigroup S and any $a, b, c \in S$, if $y \in S$ is the (b, c) -inverse of a , then y is also the Bott–Duffin (ya, ay) -inverse of a .*

Conversely, for any given idempotents $e, f \in S$, if $y \in S$ is the Bott–Duffin (e, f) -inverse of a , then y is also, in the sense of Definition 1.3, both the (e, f) -inverse of a and the (ya, ay) -inverse of a .

Proof. Given (1) and (2) for a, b, c, y , and on writing $e = ya$ and $f = ay$, then, by Theorem 2.1(ii), $e^2 = e, f^2 = f$ and we have $ey = (ya)y = y$ and dually $yf = y$, while $yae = ya(ya) = (yay)a = ya = e$ and dually $faf = f$.

Conversely, if a, e, f, y are as in Definition 3.2, then, in Definition 1.3, for the case $(b, c) = (e, f)$, we have (1) $y = ey = e^2y \in eSy$ and dually $y \in ySf$, while (2) is even more immediate. Similarly, for the case $(b, c) = (ya, ay)$, we have, by Theorem 2.1(ii), that (1) $y = yay = ya(ya)y \in yaSy$, and dually $y \in ySay$, while also (2) $ya(ya) = (yay)a = ya$ and dually $(ay)ay = ay$. \square

While the (b, c) -inverse and the Bott–Duffin (e, f) -inverse are formally very similar, the Bott–Duffin approach depends on starting from given idempotents e and f , and the essential contribution of the (b, c) -inverse lies in avoiding this. Also (even for $y = a^\dagger$ or a') one cannot use $(e, f) = (ya, ay)$ to obtain y as the Bott–Duffin (e, f) -inverse of a , since y must be known before one can find these e, f .

As regards the computation of Bott–Duffin (e, f) -inverses y in $M_n(F)$, the method outlined in Remark 2.3 still applies: y exists if and only if $e \in Sf$ and $f \in faeS$, or, for $S = M_n(F)$, equivalently $\text{rank}(f) = \text{rank}(fae) = \text{rank}(e)$, and now $y = ew$ where w is any solution of $f = faew$. This is straightforward, and works for all a -compatible pairs (e, f) , but is less efficient than the explicit Bott–Duffin definition $y = e(1 - e + ae)^{-1}$ valid when $e = f$. Thus it is worth asking whether Proposition 3.1 might be extendable even to the case of general $e \neq f$. For example, from Definition 3.2, we do have $y(1 - f + ae) = e$, so that, if $1 - f + ae$ were invertible, we could at once obtain y explicitly (the fact that a formal explicit expression for the inverse of $1 - f + ae$ may not exist, or may not be known, or may involve the unknown y , would cause no difficulty in obtaining y). However, the example

$$a = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad e = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

so that

$$1 - f + ae = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix},$$

shows that $1 - f + ae$ need not be invertible. Thus, even though $y(1 - f + ae) = e$, unfortunately $1 - f + ae$ does not usefully generalize the role of $1 - e + ae$ in Proposition 3.1.

Instead, working in rings R , it seems that we should try to construct $p, q, r \in R$, with p and r not involving y , such that $pq = qp = 1$ and $yp = r$, whence $y = rp^{-1}$. Here p and r would need to be appropriate generalizations of $p = 1 - e + ae$ and $r = e$ as in Proposition 3.1, but one might hope that at least the same $q = 1 - ay + y$ would (since it does not involve e or f) still be usable without modification. However, again unfortunately there are choices of a, e, f, y (necessarily with $e \neq f$) and also of a, b, c, y (even with $b = c$) such that $1 - ay + y$ is singular.

This is particularly easy to see for $e \neq f$, since one need consider only the Moore–Penrose situation for $a = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ with $e = a^\dagger a = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ and $f = aa^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ as after Theorem 2.1: here $y = a^\dagger = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, so that $q = 1 - ay + y = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$.

For the (b, c) -inverse, and with the same $a = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, we can find a (b, c) -inverse y of a , even with $b = c$, such that again $q = 1 - ay + y$ is singular: take $b = c = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = a^*$, so that $y = b = c$ and again $q = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$.

Given any S and $a, e, f \in S$ with e and f both idempotent, then obviously a has a Bott–Duffin (e, f) -inverse y if and only if y is a (b, c) -inverse with $b = e$ and $c = f$. Thus we can use Example 2.5 to illustrate what Definition 3.2 means in $S = M_2(F)$:

Example 3.4. If $a = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, then $b = \begin{pmatrix} \lambda \\ 1 \end{pmatrix} (\alpha, \beta)$ is idempotent if and only if $(\alpha, \beta) \begin{pmatrix} \lambda \\ 1 \end{pmatrix} = 1$, i.e. $\beta = 1 - \alpha\lambda$, so that $b = e = \begin{pmatrix} \lambda\alpha & \lambda(1 - \alpha\lambda) \\ \alpha & 1 - \alpha\lambda \end{pmatrix}$. Likewise, we need $c = f = \begin{pmatrix} 1 - \mu\delta & (1 - \mu\delta)\mu \\ \delta & \delta\mu \end{pmatrix}$, and a has Bott–Duffin (e, f) -inverse $y = \begin{pmatrix} \lambda & \lambda\mu \\ 1 & \mu \end{pmatrix}$.

Similarly, for $a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, the only possible a -compatible idempotents are

$$e = \begin{pmatrix} 1 - \beta\lambda & \beta \\ \lambda(1 - \beta\lambda) & \lambda\beta \end{pmatrix}, \quad f = \begin{pmatrix} 1 - \mu\delta & (1 - \mu\delta)\mu \\ \delta & \delta\mu \end{pmatrix},$$

and this a has Bott–Duffin (e, f) -inverse $y = \begin{pmatrix} 1 & \mu \\ \lambda & \lambda\mu \end{pmatrix}$.

Thus, much as in Example 2.5, for these a one still has ample freedom in constructing y by appropriate choice of e and f (as is also clear more generally from Proposition 3.3).

4. Extremal properties of the (b, c) -inverse

At first sight, Definition 1.3 seems to involve both the properties (1) and (2) indispensably, but, by the use of the Mitsch partial order \mathcal{M} , one can formulate two alternative descriptions of the (b, c) -inverse, respectively using either only (1) or only (2).

Definition 4.1 [15]. Given any semigroup S with 1, define a binary relation \mathcal{M} on S by saying, for any given $x, z \in S$, that $x\mathcal{M}z$ if there exist $p, q \in S$ such that

$$px = pz = x = xq = zq.$$

That \mathcal{M} is indeed a partial order on S is a routine verification. Mitsch defined \mathcal{M} so as to apply also to semigroups without unit, but we omit the details (the proofs of Lemmas 4.2 and 6.5 are valid as below even when S has no unit).

For any subset X of S , we say that X has an \mathcal{M} -maximum y if $y \in X$ and $x\mathcal{M}y$ for every $x \in X$. We write $y = \mathcal{M}\text{-max } X$. Of course X may have no \mathcal{M} -maximum in X (nor even in S), but the antisymmetry of \mathcal{M} guarantees the uniqueness of $\mathcal{M}\text{-max } X$ when it exists. We also define $\mathcal{M}\text{-min}$ similarly (the distinction between max and min amounts to interpreting $x\mathcal{M}y$ as $x \leq y$ rather than $x \geq y$).

For the sets X_a and Z_a based respectively on (1) and (2), as defined in Section 1, we have

Lemma 4.2. $x\mathcal{M}z$ for every $x \in X_a$ and every $z \in Z_a$.

Proof. If $x = xgc$, then $x = xg(caz) = (xgc)az = xaz$, and dually $x = zax$, so that

$$(xa)x = (xa)z = x = x(ax) = z(ax),$$

whence $x\mathcal{M}z$. \square

Theorem 4.3. Given any semigroup S and any $a, b, c \in S$, then

- (i) a has a (unique) (b, c) -inverse y
if and only if
- (ii) $y = \mathcal{M}\text{-max } X_a = \mathcal{M}\text{-min } Z_a$.

Proof. If (i) holds, then, by Definition 1.3 and Theorem 2.1(ii), we have $y \in X_a \cap Z_a$, and so, by Lemma 4.2, $x\mathcal{M}y$ for all $x \in X_a$ and $y\mathcal{M}z$ for all $z \in Z_a$, which is (ii).

Conversely, if (ii) holds, then $y \in X_a \cap Z_a$, and so y satisfies Definition 1.3, whence (i) follows by Theorem 2.1(i). \square

Thus, as alternatives to Definition 1.3 and Theorem 2.1(i), we have the choice to use instead either $\mathcal{M}\text{-max } X_a$ alone or $\mathcal{M}\text{-min } Z_a$ alone to define the (b, c) -inverse. Of these two, $\mathcal{M}\text{-min } Z_a$ seems to be the less satisfactory, since it can make sense only when Z_a is non-empty (while X_a has the advantage of being self-evidently non-empty whenever a zero element is available).

Example 4.4. Let $S = \mathbb{Z}$, and let $a, b, c \in \mathbb{Z}$ with $a, b, c \notin \{0, \pm 1\}$. Then $b \notin Scab$ and $c \notin cabS$, so that a has no (b, c) -inverse (and Z_a is empty). However, $X_a = \{0\}$, and so $\mathcal{M}\text{-max } X_a$ exists.

Thus $\mathcal{M}\text{-max } X_a$ may have a well-defined meaning even when no $y \in S$ satisfies Definition 1.3. Accordingly, whenever $y = \mathcal{M}\text{-max } X_a$ exists, we may call it the *extended* (b, c) -inverse of a .

The discussion above is in the spirit of [8] (see also [9]), which was focused specifically on a^\dagger and a' rather than on the more general Definition 1.3, and also used the fact that (with appropriately differently defined X_a and Z_a) if Z_a is non-empty then so is $X_a \cap Z_a$. However, I have not been able to establish this for our present X_a and Z_a .

5. Weak invertibility

Because of the possibility that $b = c = 0$, the mere existence of some a -compatible pair (b, c) tells us nothing about a . However, at least for rings (always taken to be associative with 1), the weakness of the conditions (1) and (2) allows us some freedom to add further (consistent) requirements so as to usefully restrict (b, c) :

Definition 5.1. Given any ring R with 1, we call $a \in R$ *weakly invertible* if there exist $b, c, y \in R$ such that (1) and (2) hold and also

- (3) $1 - b \in (1 - a)R$ and $1 - c \in R(1 - a)$,
- (4) $ay = ya$.

(One can also correspondingly define extended weak invertibility, much as in Section 4, but this does not seem to be helpful in our present application.)

If a is pseudo-invertible (or equivalently strongly π -regular, i.e. $a^j \in (a^{j+1}R) \cap (Ra^{j+1})$ for some $j \in \mathbb{N}$), then b, c, y satisfying Definition 5.1 always exist, with $b = c = a^j$ as before and $y = a'$, since then (1) and (2) hold as noted in the Introduction, while (3) holds because

$$(1 - a)(1 + a + \cdots + a^{j-1}) = 1 - a^j = (1 + a + \cdots + a^{j-1})(1 - a),$$

and (4) is part of the definition of $y = a'$. (Alternatively, as the referee pointed out, one may take $b = c = a'a$.)

On the other hand, for the Moore–Penrose inverse $y = a^\dagger$ (with $b = c = a^*$ as before), of course (4) fails, and so in fact does (3): consider e.g. $a = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ in $S = M_2(\mathbb{C})$. Similarly, obviously even

the original Bott–Duffin inverse does not satisfy (3): take $a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $e = y = 0$, so that $e^* = e$, $ae + 1 - e$ is invertible and (with $b = c = e = 0$) (3) fails.

For further insight into (3) and (4) one can refer to Example 2.4, where, for $a = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, $1 - a$ is invertible, so that (3) is true for all b, c , while, since every nonzero y has the form $y = \begin{pmatrix} \lambda & \lambda\mu \\ 1 & \mu \end{pmatrix}$,

(4) is always false except when $y = 0$. Similarly, for $a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, we find that, for (3) to hold, the

only possibilities for (b, c) are with $b = \begin{pmatrix} 1 & 0 \\ \lambda & 0 \end{pmatrix}$, $c = \begin{pmatrix} 1 & \mu \\ 0 & 0 \end{pmatrix}$ (in particular, b and c must both be

idempotent), while also, whether or not (3) holds, (4) is true if and only if $b = \begin{pmatrix} \alpha & \beta \\ 0 & 0 \end{pmatrix}$, $c = \begin{pmatrix} \gamma & 0 \\ \delta & 0 \end{pmatrix}$, $y = a$, so that (3) and (4) can hold simultaneously only if $b = c = y = a$. Thus, while all of (1), \dots , (4) are true for the pseudo-inverse a' , in combination they seem to be quite strong.

For given a , while (by Theorem 2.1(i)) y in Definition 5.1 is uniquely determined by any fixed (a -compatible) choice of b, c , different choices of b and c may yield different values of y . For example, if R is a field (or division ring) and $a \neq 0, 1$ (and $R \neq GF(2)$), then $b = c = 0$ gives $y = 0$, while $b = c = 1$ (or $b = c = a^{-1}$) gives $y = a^{-1}$. Nevertheless, even though it does not usually provide a single unique y corresponding to a , weak invertibility has quite strong consequences. Generalizing a result of Rothblum [19, p. 646, Theorem 1] for the case $R = M_n(\mathbb{C})$ with $b = c = a^j$ in (i), we have:

Lemma 5.2. *Given any associative ring R with 1 and any $a, y \in R$, then $u = a - 1 + ay$ is a two-sided unit of R in each of the following three cases:*

- (i) a is weakly invertible, with y (and some $b, c \in R$) as in Definition 5.1;
- (ii) $yay = y$, $1 - y \in ((1 - a)R) \cap (R(1 - a))$ and $ay = ya$;
- (iii) $yay = y$, $1 - ay \in ((1 - a)(1 - ay)R) \cap (R(1 - ay)(1 - a))$ and $ay = ya$.

Proof. Set $e = 1 - ay = 1 - ya$, so that (by Theorem 2.1(ii) in case (i)) $e^2 = e$, $ae = ea$, $ey = 0$ and $u = a - e$.

(i) By (3) we have $1 - b = (1 - a)s$ and $1 - c = t(1 - a)$ for suitable $s, t \in R$, while (2) gives $eb = 0$. Hence

$$\begin{aligned} u(y - es) &= (a - e)(y - es) = ay - ey - aes + es \\ &= ay + e(1 - a)s = ay + e(1 - b) = 1, \end{aligned}$$

and dually $(y - te)u = 1$.

(ii) By Theorem 2.1(ii), this is just the special case of (i) with $b = c = y$ (since (1) and (2) both hold trivially for this choice of b, c).

(iii) We are given that $e \in ((1 - a)eR) \cap (Re(1 - a))$, say with $e = (1 - a)es = te(1 - a)$. Hence

$$u(y - es) = ay - ey - aes + es = ay + (1 - a)es = ay + e = 1,$$

and dually $(y - te)u = 1$. \square

Of course the hypothesis that $yay = y$ is crucial in each case; I have not been able to find any version of Lemma 5.2 without the assumption that $ay = ya$.

Corollary 5.3. *In each of the three cases of Lemma 5.2, we have $(1 - y)R = (1 - a)R$ and $R(1 - y) = R(1 - a)$.*

Proof. $(1 - y)u = a - 1 + ay - ya + y - yay = a - 1$, whence $(1 - y)R = (1 - a)R$, and dually $R(1 - y) = R(1 - a)$. \square

It is surprising that (i), (ii) and (iii) of Lemma 5.2 should each imply that $1 - a \in ((1 - y)R) \cap (R(1 - y))$.

Remark 5.4. Somewhat analogously to Lemma 5.2, but even without needing (1), (2) or (3), if we assume that $a, y \in R$ satisfy just $yay = y$ and $ay = ya$, then, by taking $e = ay = ya$ in Proposition 3.1, we have that $1 - ay + y$ is a unit, with inverse $1 - ay + aya$ (whence, as before, both y and aya are unit regular).

Indeed, here we can even relax the requirement that $ay = ya$. For example, if, while still assuming that $yay = y$, we replace $ay = ya$ by the weaker requirement that $ay^2 = y$ and $ya^2y = ay$, then it is easy to verify that

$$(1 - ya + y)(1 - ay + aya) = 1 = (1 - ay + aya)(1 - ya + y).$$

To confirm that the hypotheses $yay = y$, $ay^2 = y$ and $ya^2y = ay$ are, collectively, strictly weaker than $yay = y$ with $ay = ya$, it suffices to consider the example $a = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, y = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ in $M_2(F)$.

Even assuming only $yay = y$ without any kind of commutativity, non-trivial “trinomial” units do still exist: given any $a, y, z \in R$ with $yay = y$, then $n = yz(1 - ya)$ obviously satisfies $n^2 = 0$, so that $(1 + n)(1 - n) = (1 - n)(1 + n) = 1$. I am indebted to Studzinski, who, at my request and using code (see [11–14]) created and developed by herself and Levandovskyy et al., found the special case $z = y^j$ (I also thank Uli Walther for telling me about this code). However, I have not yet been able to find a choice of z having any significant application.

Nicholson [16] has defined an element $a \in R$ to be *clean* if a is the sum $a = u + e$ of some unit u and an idempotent e . He calls a *strongly clean* [17] if also $eu = ue$, and this property is equivalent to weak invertibility etc.:

Theorem 5.5. *Let R be any associative ring with 1. Then, for any fixed $a \in R$, if there exists $y \in R$ satisfying (i), (ii) or (iii) of Lemma 5.2, then a is strongly clean.*

Conversely, if $a \in R$ is strongly clean, then there exists $y \in R$ which simultaneously satisfies all of (i), (ii) and (iii).

Proof. By Lemma 5.2, if any of (i), (ii) or (iii) holds, then $a = u + (1 - ay) = u + e$, where u is a unit and $e = 1 - ay$ is idempotent, while $eu = ue$ since $ay = ya$.

Conversely, if $a = u + e$ with $eu = ue$, then, on regarding u and e as fixed, and on taking $x = u^{-1}(1 - e) = (1 - e)u^{-1}$, we have

$$\begin{aligned} ax &= (u + e)(1 - e)u^{-1} = u(1 - e)u^{-1} = 1 - e, \\ xax &= u^{-1}(1 - e)(1 - e) = u^{-1}(1 - e) = x, \\ 1 - x &= 1 - (1 - e)u^{-1} = (u - 1 + e)u^{-1} = (a - 1)u^{-1} \in (1 - a)R, \end{aligned}$$

and dually $1 - x \in R(1 - a)$, while also $ae = ea$ and $au = ua$, so that $ax = xa$. Thus x satisfies (ii), hence also (i) (on taking $b = c = x$ as above).

Finally, for (iii), with the same x , we have $eu = e(a - e) = e(a - 1) = (a - 1)e$, so that $e = (1 - a)e(-u^{-1}) \in (1 - a)eR$, and dually $e \in Re(1 - a)$, which gives (iii). \square

Corollary 5.6. For any associative ring R with 1 and any fixed $a \in R$, there is a natural bijective correspondence between the set of all $y \in R$ satisfying (i) (or (ii), or (iii)) and the set of all strongly clean decompositions $a = u + e$.

Proof. For fixed $a \in R$, let $G = G_a$ denote the set of all $y \in R$ satisfying (i) [resp. (ii) or (iii)] of Lemma 5.2, and let $H = H_a$ denote the set of all pairs $(u, e) \in R \times R$ for which $a = u + e$ is a strongly clean decomposition of a . By Lemma 5.2, the assignment $\phi(y) = (a - 1 + ay, 1 - ay)$ defines a map $\phi: G \rightarrow H$.

To prove that ϕ is injective, choose any fixed $(u, e) \in H$, and suppose that $\phi(y) = (u, e)$ for some $y \in G$, i.e. that $u = a - 1 + ay$ and $e = 1 - ay = 1 - ya$. Then we must have

$$y = yay = (1 - e)y = (1 - e)(y - es) = (1 - e)u^{-1}$$

by the proof of Lemma 5.2, so that a given pair $(u, e) \in H$ can have at most one pre-image under ϕ , namely $(1 - e)u^{-1}$.

To prove that ϕ is also surjective, it now remains only to show, for any fixed $(u, e) \in H$, that indeed $x = (1 - e)u^{-1} \in G$, which holds as in the proof of Theorem 5.5, and, finally, that $\phi(x) = (u, e)$. But, since (as above) $ax = 1 - e$, we have at once that $a - 1 + ax = a - e = u$ and $1 - ax = e$, i.e. $\phi(x) = (u, e)$, as required. \square

By Theorem 5.5 and what is already known about clean rings, weak invertibility also implies (e.g.) the exchange property, suitability and potency (see [16, pp. 271–274]), which includes the definitions). One may also consider defining a new property (or properties) stronger than weak invertibility by adding further requirements to Definition 5.1. For example, to obtain a property strong enough to imply direct finiteness (i.e. $ar = 1 \Rightarrow ra = 1$), it suffices to add to Definition 5.1 the requirement

$$(*) \quad aR = R \Rightarrow bR = R$$

(obviously true when $b = a^j$), since, if $ar = 1$ and $bd = 1$, we then have, for any y satisfying (2), that

$$y = y(a(bd)r) = (yab)dr = b(dr) = (bd)r = r,$$

so that, by (4), $ra = ya = ay = ar = 1$. However, $(*)$ seems unsatisfying, and one might hope to find a better substitute for it.

An ultimate objective in this direction would be to formulate a stronger version of Definition 5.1 (still implied by strong π -regularity) which implies not only direct finiteness but also the stable range 1 property of Bass [2], so as to strengthen Ara's remarkable result [1] that strong π -regularity implies stable range 1. Possibly something like Lemma 5.2 (or Remark 5.4) could be useful for this.

6. Some variations on the theme

In Theorem 2.1(i), in order to prove the uniqueness of y without assuming that $yay = y$, we needed to use the full force of (1) and (2). In this section we explore how far (1) and (2) can be relaxed if we choose instead to assume explicitly from the start that $yay = y$.

Proposition 6.1. *Given any semigroup S with 1 and any $a, b, c, y \in S$ such that $yay = y$, then (1) holds if and only if $y \in (bS) \cap (Sc)$, and (2) holds if and only if $b \in yS$ and $c \in Sy$.*

Proof. If (1) holds then $y \in bSy \subseteq bS$, while conversely $y \in bS$ gives $y = yay \in (bS)ay \subseteq bSy$; and dually $y \in ySc$ if and only if $y \in Sc$.

If (2) holds, then $b = yab \in yS$, while conversely, if $b \in yS$, say with $b = ys$ for suitable $s \in S$, then $b = (yay)s = ya(ys) = yab$; and dually $cay = c$ if and only if $c \in Sy$. \square

By Theorem 2.1(ii), it follows that (1) and (2) combined are equivalent to

$$(5) \quad yay = y, \quad yS = bS \text{ and } Sy = Sc.$$

An advantage of using (5) rather than (1) and (2) is that (5) immediately suggests two other closely related but somewhat weaker conditions on a, b, c, y . These could be straightforwardly explored even for semigroups, but for simplicity we shall work here only with rings R (however, when $R = M_n(F)$, our alternative versions (6) and (7) below become equivalent to (5)).

Definition 6.2. For any associative ring R with 1 and any $a, b, c, y \in R$, we call y an *annihilator* (b, c) -inverse of a if

$$(6) \quad yay = y, \quad \text{lann}(y) = \text{lann}(b) \text{ and } \text{rann}(y) = \text{rann}(c).$$

Definition 6.3. With R, a, b, c, y as in Definition 6.2, call y a *hybrid* (b, c) -inverse of a if

$$(7) \quad yay = y, \quad yR = bR \text{ and } \text{rann}(y) = \text{rann}(c).$$

For matrices, this hybrid property (with $b = c$) was discussed at length by Ben-Israel and Greville [3, p. 72, Theorem 14], and since then has been studied by many other writers.

Obviously (5) \Rightarrow (7) \Rightarrow (6), and we next revisit our results of Sections 2 and 4 to find which can be strengthened by using Definition 6.2 or 6.3 instead of Definition 1.3.

In regard to Theorem 2.2, Roberts and I ([10], Theorems 1.3 and 2.2, also Lemma 2.1) have shown that, given $a, b, c \in R$, then a hybrid (b, c) -inverse y of a exists if and only if $\text{rann}(cab) \subseteq \text{rann}(b)$ and $c \in cabR$. However, while it is easy to see that (6) implies $\text{rann}(cab) \subseteq \text{rann}(b)$ and $\text{lann}(cab) \subseteq \text{lann}(c)$ (or equivalently $\text{rann}(cab) = \text{rann}(b)$ and $\text{lann}(cab) = \text{lann}(c)$), I have not been able to find any useful condition necessary and sufficient for the existence of annihilator (b, c) -inverses as in (6).

The situation regarding analogues of Theorem 2.1(i) is more satisfactory, in that the uniqueness of y follows just as easily from (6) or (7) as it does from (1) and (2) together:

Theorem 6.4. *For any associative ring R with 1 and any $a, b, c \in R$, there can be at most one annihilator [resp. hybrid] (b, c) -inverse y of a .*

Proof. Since (7) \Rightarrow (6), it suffices to prove this for the annihilator (b, c) -inverse. So suppose that x and $y \in R$ both satisfy (6). Then $1 - xa \in \text{lann}(x) = \text{lann}(b) = \text{lann}(y)$, i.e. $xay = y$, and dually $1 - ay \in \text{rann}(y) = \text{rann}(c) = \text{rann}(x)$ gives $xay = x$. Hence $x = xay = y$. \square

Turning next to possible annihilator or hybrid analogues of our results in Section 4, by Proposition 6.1 our X_a and Z_a of Sections 1 and 4 can equivalently be written as

$$X_a = \{x: x \in S, xax = x \text{ and } x \in (bS) \cap (Sc)\},$$

$$Z_a = \{z: z \in S, zaz = z, b \in zS \text{ and } c \in Sz\},$$

and we must now consider instead

$$\text{ann-}X_a = \{x: x \in R, xax = x, \text{lann}(b) \subseteq \text{lann}(x) \text{ and } \text{rann}(c) \subseteq \text{rann}(x)\},$$

$$\text{ann-}Z_a = \{z: z \in R, zaz = z, \text{lann}(z) \subseteq \text{lann}(b) \text{ and } \text{rann}(z) \subseteq \text{rann}(c)\},$$

$$\text{hy-}X_a = \{x: x \in R, xax = x, x \in bR \text{ and } \text{rann}(c) \subseteq \text{rann}(x)\},$$

$$\text{hy-}Z_a = \{z: z \in R, zaz = z, b \in zR \text{ and } \text{rann}(z) \subseteq \text{rann}(c)\},$$

where obviously $X_a \subseteq \text{hy-}X_a \subseteq \text{ann-}X_a$ and $Z_a \subseteq \text{hy-}Z_a \subseteq \text{ann-}Z_a$.

Again, Lemma 4.2 extends immediately to both the annihilator and the hybrid contexts:

Lemma 6.5. $x\mathcal{M}z$ for every $x \in \text{ann-}X_a$ and every $z \in \text{ann-}Z_a$.

Proof. Let $x \in \text{ann-}X_a$ and $z \in \text{ann-}Z_a$. Then $1 - za \in \text{lann}(z) \subseteq \text{lann}(b) \subseteq \text{lann}(x)$, i.e. $zax = x$, and dually $xaz = x$. Thus $x\mathcal{M}z$ as in Lemma 4.2. \square

By exactly the same argument as for Theorem 4.3, we obtain

Theorem 6.6. Given any associative ring R with 1 and any $a, b, c \in R$, then a has a (unique) annihilator [resp. hybrid] (b, c) -inverse y if and only if

$$y = \mathcal{M}\text{-max}(\text{ann-}X_a) = \mathcal{M}\text{-min}(\text{ann-}Z_a)$$

$$[\text{resp } y = \mathcal{M}\text{-max}(\text{hy-}X_a) = \mathcal{M}\text{-min}(\text{hy-}Z_a)]. \quad \square$$

References

- [1] P. Ara, Strongly π -regular rings have stable range one, Proc. Amer. Math. Soc. 124 (1996) 3293–3298.
- [2] H. Bass, K -theory and stable algebra, Inst. Hautes Études Sci. Publ. Math. 22 (1964) 5–60.
- [3] A. Ben-Israel, T.N.E. Greville, Generalized Inverses: Theory and Applications, second ed., Springer, NY, 2003.
- [4] R. Bott, R.J. Duffin, On the algebra of networks, Trans. Amer. Math. Soc. 74 (1953) 99–109.
- [5] T.L. Boullion, P.L. Odell (Eds.), Proceedings of the Symposium on Theory and Applications of Generalized Inverses, Texas Tech Press, Lubbock, 1968.
- [6] Randall E. Cline, T.N.E. Greville, A Drazin inverse for rectangular matrices, Linear Algebra Appl. 29 (1980) 53–62.
- [7] M.P. Drazin, Pseudo-inverses in associative rings and semigroups, Amer. Math. Monthly 65 (1958) 506–514.
- [8] M.P. Drazin, Extremal definitions of generalized inverses, Linear Algebra Appl. 165 (1992) 185–196.
- [9] M.P. Drazin, Generalizations of Fitting's lemma in arbitrary associative rings, Comm. Algebra 29 (2001) 3647–3675.
- [10] M.P. Drazin, M.L. Roberts, Some finiteness conditions for rings or semigroups, in press.
- [11] V. Levandovskyy, H. Schönemann; with cooperation by G. Studzinski and B. Schnitzler, Singular:Letterplace – a Singular 3-1-3 subsystem for computations with free associative algebras, 2011. <<http://www.singular.uni-kl.de>>.
- [12] V. Levandovskyy, G. Studzinski, A Singular 3.1 library for computing two-sided Groebner bases in free algebras via letterplace methods, 2011. <<http://www.singular.uni-kl.de>>.
- [13] V. Levandovskyy, H. Schönemann, Singular:Letterplace – a Singular 3-1-3 subsystem for computations with free associative algebras, 2011. <<http://www.singular.uni-kl.de>>.
- [14] V. Levandovskyy, G. Studzinski, A Singular 3.1 library for computing two-sided Groebner bases in free algebras via letterplace methods, 2011. <<http://www.singular.uni-kl.de>>.
- [15] H. Mitsch, A natural partial order for semigroups, Proc. Amer. Math. Soc. 97 (1986) 384–388.
- [16] W.K. Nicholson, Lifting idempotents and exchange rings, Trans. Amer. Math. Soc. 229 (1977) 269–278.
- [17] W.K. Nicholson, Strongly clean rings and Fitting's lemma, Comm. Algebra 27 (1999) 3583–3592.
- [18] R. Penrose, A generalized inverse for matrices, Proc. Cambridge Philos. Soc. 51 (1955) 406–413.
- [19] U.G. Rothblum, A representation of the Drazin inverse and characterizations of the index, SIAM J. Appl. Math. 31 (1976) 646–648.