

## CONTRIBUTIONS TO THE THEORY OF GENERALIZED INVERSES\* †

A. BEN-ISRAEL‡ AND A. CHARNES§

### INTRODUCTION<sup>1</sup>

In this introduction we review briefly some developments in the theory of generalized inverses of linear operators on a Euclidean or Hilbert space, and the underlying theory of regular rings.

The spaces considered are:

(i). *Finite-dimensional vector spaces over the complex field*, to be denoted by  $E^n$ ,  $E^m$ .

(ii). *Complex Hilbert spaces, separable or not*, to be denoted by  $\mathfrak{H}^1$ ,  $\mathfrak{H}^2$ .

Accordingly, the linear operators considered are:

(i). Linear transformations  $E^n \rightarrow E^m$  represented as *complex matrices*, denoted by  $A$ .

(ii). *Closed linear operators*  $A: \mathfrak{H}^1 \rightarrow \mathfrak{H}^2$  whose domain is dense in  $\mathfrak{H}^1$ .

With  $A$  as a generic notation for the linear operator considered, we denote the *domain*, *range*, and *null-manifold* of  $A$  by  $D(A)$ ,  $R(A)$  and  $N(A)$  respectively. By  $A^*$  we denote the conjugate of  $A$  (*conjugate transpose* for a matrix  $A$ ).

In  $E^n$  we use the Euclidean norm

$$\|x\| = \left( \sum_{i=1}^n x_i \bar{x}_i \right)^{\frac{1}{2}}$$

and a compatible matrix norm,<sup>3</sup>

$$\|A\| = \max_{i=1, \dots, n} \{ \sqrt{\lambda_i} : \lambda_i \in \sigma(A^*A) \},$$

where  $\sigma(A)$  denotes the *spectrum* of  $A$ .

\* Received by the editors May 10, 1962, and in final revised form January 14, 1963. The research underlying this paper was supported in part by the U. S. Office of Naval Research (Contract Nonr-1228(10), Project NR 047-021) and in part by the National Science Foundation (Contract NSF-614102). The contribution of A. Ben-Israel constitutes part of his dissertation for the Ph.D. degree at Northwestern University.

† The authors are greatly indebted to the referees for many useful remarks especially on Part II.

‡ Carnegie Institute of Technology, Pittsburgh 13, Pennsylvania.

§ The Technological Institute, Northwestern University, Evanston, Illinois.

<sup>1</sup> We are indebted to L. Martić of the University of Zagreb, for his careful translation of [48], which is incorporated into § 2 of this introduction.

<sup>2</sup> The restriction to complex Hilbert spaces is not essential; e.g., Tseng's results hold in Hilbert spaces over the quaternions as well.

<sup>3</sup> See [16, pp. 55-60].

In  $\mathcal{H}$  we denote the *closure* of a set  $S$  by  $\bar{S}$ .

The *projection on a subspace*  $L$  will be denoted by  $P_L$ , the *orthogonal complement* of a set  $S$  by  $O(S)$  and for two subspaces  $L, M$ , we denote the *direct sum* (in case  $L \subseteq O(M)$ ) by  $L \oplus M$  and the *difference* (in case  $M \subseteq L$ ) by  $L \ominus M$ .

The *generalized inverse*<sup>4</sup> of  $A$ , denoted by  $A^+$  and abbreviated g.i. will be defined below, and the main results and references will be presented.

**1. Matrices.** The concept of g.i. for arbitrary  $m \times n$  matrices with components from the complex field is due to Moore [29, 30], whose definition is essentially:

*Definition 1.*  $A^+$  is the g.i. of  $A$  if<sup>5</sup>

$$(1) \quad AA^+ = P_{R(A)},$$

$$(2) \quad A^+A = P_{R(A^+)}.$$

Moore established the existence and uniqueness of  $A^+$  for any  $A$ , and gave an explicit form for  $A^+$  in terms of the subdeterminants of  $A$  and  $A^*$ . Moore's various results on  $A^+$  and the relations between  $A, A^*$  and  $A^+$  were incorporated in his *General Analysis*, and concurrently were given an algebraic basis and extensions in Von Neumann's studies on regular rings (see §3).

In extending Moore's results to closed linear operators on a Hilbert space, Tseng investigated in [45, 48] virtual solutions of linear operator equations. His results (e.g. Theorem 7 below) suggest for matrices the following definition (equivalent to definition 1) which rests on the *least-square* character of solutions to linear equations, obtained by using  $A^+$ .

*Definition*<sup>6</sup> 2. Consider the equation

$$(3) \quad Ax = a, \quad a \in E^m, x \text{ unknown} \in E^n.$$

Among all *virtual solutions*  $x_a$  of (3), defined as

$$(4) \quad \|Ax_a - a\| = \inf_{x \in E^n} \|Ax - a\|,$$

there is a unique *extremal virtual solution*  $x_a^0$  defined as

$$(5) \quad \|x_a^0\| = \inf \{\|x_a\| : x_a \text{ satisfies (4)}\}.$$

<sup>4</sup> Following Penrose we use this name rather than *pseudo inverse*, used by Drazin and Greville (introduced by Fredholm in a different context) or *general reciprocal* used by Moore, Wong and Hestenes. For generalized inverses of matrices we will occasionally use the name *Moore-Penrose inverses*. Similarly, g.i. of closed operators on a Hilbert space will be called here *Moore-Tseng inverses*.

<sup>5</sup> In what follows we assume implicitly the definability of matrix products, thus  $A^+$  is an  $n \times m$  matrix.

<sup>6</sup> This definition was suggested by Penrose in [35, p. 18].

The g.i.  $A^+$  of  $A$  is the matrix corresponding to the linear transformation  $a \rightarrow x_a^0$  as  $a$  varies in  $E^m$ .

Unaware of Moore's work<sup>7</sup>, the g.i. was treated independently by Bjerhammar [5, 6] and Penrose [34]. Bjerhammar constructed  $A^+$  by identifying it with a submatrix of the inverse of a suitable square nonsingular matrix, obtained by multiplying  $A$  with another matrix. The general solution of (3), when solvable, was given by Bjerhammar as

$$(6) \quad x = A^+a + (I - A^+A)y, \quad y \text{ arbitrary } \in E^n$$

which is a corollary of Theorem 1 below. The least square character of the solution  $A^+a$  was used by Bjerhammar in geodetic applications: adjusting observations which gave rise to singular or ill-conditioned matrices.

Penrose in [34] defined the g.i. as follows (clearly equivalent to definition 1).

*Definition 3.*  $A^+$  is a solution of

$$(7) \quad AXA = A,$$

$$(8) \quad XAX = X,$$

$$(9) \quad (AX)^* = AX,$$

$$(10) \quad (XA)^* = XA.$$

Penrose's proof of the existence and uniqueness of  $A^+$  is based on the vanishing of a finite polynomial in<sup>8</sup>  $A^*A$ . Some of his other results are summarized below:

LEMMA 1. (*Properties of  $A^+$* ).

$$(a). \quad A^{++} = A.$$

$$(b). \quad A^{**} = A^{**}.$$

$$(c). \quad |A| \neq 0 \Rightarrow A^+ = A^{-1}.$$

$$(d). \quad (\lambda A)^+ = \lambda^+ A^+, \text{ where for a scalar } \lambda, \lambda^+ = \begin{cases} \frac{1}{\lambda} & \text{if } \lambda \neq 0, \\ 0 & \text{if } \lambda = 0. \end{cases}$$

$$(e). \quad (A^*A)^+ = A^+A^{**}.$$

$$(f). \quad \text{If } A = \sum_i A_i \text{ and } i \neq j \Rightarrow A_i A_j^* = 0 = A_i^* A_j, \text{ then } A^+ = \sum_i A_i^+.$$

$$(g). \quad A \text{ normal} \Rightarrow A^+A = AA^+, (A^n)^+ = (A^+)^n.$$

$$(h). \quad \text{The ranks of } A, A^*A, A^+, A^+A \text{ are all equal to the trace of } A^+A.$$

<sup>7</sup> Not too well known, because the unique notations employed there were not adopted by other mathematicians. An outstanding account of some of Moore's results is given by Greville [19] where the theory is elegantly redeveloped.

<sup>8</sup> A proof valid for infinite matrices as well is given by Ben-Israel and Wersan [4].

<sup>9</sup> Note that generally  $(AB)^+ \neq B^+A^+$ .

**THEOREM 1.** (*Solvability of linear equations.*) *A necessary and sufficient condition for the solvability of*

$$(11) \quad AXB = C$$

*is that*

$$(12) \quad AA^+CB^+B = C,$$

*in which case the general solution of (11) is*

$$(13) \quad X = A^+CB^+ + (Y - A^+AYBB^+)$$

*where  $Y$  is an arbitrary matrix, of the same size as<sup>10</sup>  $X$ .*

**THEOREM 2.** (*Explicit form for the principal idempotents.*) *The principal idempotents  $E_\lambda$  of<sup>11</sup>  $A$  are given by*

$$(14) \quad E_\lambda = (F_\lambda G_\lambda)^+$$

*where*

$$(15) \quad F_\lambda = I - (A - \lambda I)^n \{(A - \lambda I)^n\}^+,$$

$$(16) \quad G_\lambda = I - \{(A - \lambda I)^n\}^+ (A - \lambda I)^n$$

*and  $n$  is a positive integer depending on  $A$ , which can be taken as  $n = 1$  if and only if  $A$  is diagonalable.*

**THEOREM 3.** (*Spectral decomposition.*)<sup>12</sup> *Any matrix  $A$  may be uniquely represented as a linear combination of partial isometries*

$$(17) \quad A = \sum_i \alpha_i U_{\alpha_i},$$

*where  $\alpha_i \in \sigma(A^*A)$  and the matrices*

$$(18) \quad U_{\alpha_i} = \alpha_i^+ A \{I - (A^*A - \alpha_i^2 I)^+ (A^*A - \alpha_i^2 I)\}$$

*satisfy*

$$(19) \quad \alpha_i \neq \alpha_j \Rightarrow U_{\alpha_i}^* U_{\alpha_j} = 0 = U_{\alpha_i} U_{\alpha_j}^*.$$

*The g.i.  $A^+$  is given by*

$$(20) \quad A^+ = \sum_i \alpha_i^+ U_{\alpha_i}.$$

<sup>10</sup> Theorem 1 is valid for any  $A^+$ ,  $B^+$  which respectively satisfy  $AA^+A = A$ ,  $BB^+B = B$ .

<sup>11</sup> See also Wedderburn [53, p. 29].

<sup>12</sup> Hestenes pointed out in [21, p. 88] that this decomposition and other related results are due to Gibbs, see [54].

Corollary 1. (Polar representation.)<sup>13</sup> Any matrix  $A$  is uniquely represented as

$$(21) \quad A = BW$$

where  $B = \sqrt{AA^*}$  is Hermitian nonnegative definite, and  $W$  is a partial isometry (i.e.  $W^* = W^+$ ) such that<sup>14</sup>  $WW^* = AA^+$ .

Penrose suggested in [35] applications of the g.i. in least square solutions to inconsistent linear equations, in particular to statistical problems, and gave two methods to calculate  $A^+$ . One method is based on a suitable partition of  $A$ , allowing an expression of  $A^+$  in terms of the (regular) inverses of the partitioned submatrices. The second method is an iterative procedure, involving the subdeterminants of  $A^*A$ .

Rado [37] extended Moore's results to matrices over any division ring with an involutory anti-automorphism

$$x \rightarrow \bar{x} \quad \text{such that} \quad \sum_i x_i \bar{x}_i = 0 \Rightarrow x_i = 0.$$

Explicit expressions for  $A^+$  as a limit, were given by den Broeder and Charnes [11]. Their following two theorems are based on Autonne's theorem that any square matrix  $A$  can be represented as<sup>15</sup>

$$(22) \quad A = VDW,$$

$D$  diagonal,  $V$  and  $W$  unitary.

THEOREM<sup>16</sup> 4. For any square matrix  $A$ ,  $\lim_{n \rightarrow \infty} \sum_{k=1}^n A^*(I + AA^*)^{-k}$  exists, and

$$(23) \quad A^+ = \sum_{k=1}^{\infty} A^*(I + AA^*)^{-k}$$

(where  $A^*$  may not be removed as a factor from the series).

THEOREM 5. For any square matrix  $A$ ,

$$(24) \quad A^+ = \lim_{\lambda \rightarrow 0} A^*(\lambda \bar{\lambda} I + AA^*)^{-1}.$$

The other results by den Broeder and Charnes include some theorems on the g.i., rank and conditions on nonsingularity for some matrices of

<sup>13</sup> See Von Neumann [49], p. 307 theorem 7.

<sup>14</sup>  $W$ , so normalized, is unique.

<sup>15</sup> As noted by Penrose [34, p. 908], this suggests the "constructive" definition  $A^+ = W^*D^+V^*$  where  $D^+ = (d_{ii}^+)$ .

<sup>16</sup> Without loss of generality we may consider only square matrices  $A$ , as any  $m \times n$  matrix  $B$  may be written as a square matrix  $A$  by adding a proper number of zero columns or rows. The  $n \times m$  g.i.  $B^+$  is then identified with the corresponding submatrix in  $A^+$ .

special structure, and a necessary and sufficient condition for  $A$  to be a solution of the circle composition equation<sup>17</sup>

$$AX = A + X = XA$$

where  $A$  is normal.

They applied the form of  $A^+$  in Theorem 5 to solve a problem in diffusion. Greville [19] gave a very clear and suggestive exposition of g.i. of matrices following the original Moore approach.

A review of the various definitions and applications in explicit solutions of systems of linear equations was given by Bjerhammar [7] together with numerical examples and statistical applications.

Hestenes [21] gave a method for inverting nonsingular matrices by reducing "inversion" to the equivalent problem of constructing suitable biorthogonal systems of vectors. In carrying his results to the general case, he extended Autonne's theorem (22) to rectangular matrices, where instead of unitary matrices,  $V$  and  $W$  (in (22)) are partial isometries given in terms of the "principal vectors" of  $A, A^*$ . In sharpening the Gibbs-Penrose decomposition (17), Hestenes characterized the set of matrices  $\mathfrak{B}$  of the form

$$B_f = f(A) = \sum_i f(\alpha_i) U_{\alpha_i}, \quad f(x) \text{ real function,}$$

as all the matrices  $B$  satisfying for  $U = \sum_i U_{\alpha_i}$ ,

$$BU^*C = CU^*B,$$

for all the matrices  $C$  satisfying

$$CU^*A = AU^*C, \quad P_{R(A)}C = CP_{R(A^+)} = C.$$

Developing a spectral theory for arbitrary  $m \times n$  matrices, which is an extension of the Hermitian case theory, Hestenes [22] used  $A^+$  in an essential manner to obtain theorems on structure and some properties of matrices relative to "elementary matrices" and the relations of: "\*-orthogonality" and "\*-commutativity." His main results can be extended to the case of a closed linear operator between two Hilbert spaces [22, p. 225].

Greville [20] gave an iterative procedure for calculating  $A^+$  using successive partitions of  $A$ . Using  $A^+$  he modified the procedure of Dent and Newhouse [12] in constructing polynomials orthogonal over a discrete domain, and used the least square properties of  $A^+$  in regression analysis.

<sup>17</sup> Similar to a *quasi-inverse* in the ring-theoretic sense:  $(A - I)(X - I) = (X - I)(A - I) = I$ .

Pyle [36] and Cline [10] following on den Broeder and Charnes have considered applications to systems of linear equations. The projections  $AA^+$ ,  $A^+A$  were used by Pyle [36] in a gradient method of solving linear programming problems. These were also used by Rosen [41, 42] in his conjugate gradient method of solving linear and nonlinear programs.

Recently, Charnes, Cooper and Thompson have employed g.i. and the associated solvability criteria in an essential manner to resolve questions of the scope and validity of so called "linear programming under uncertainty" and to characterize optimal stochastic decision rules. Kalman [24, 25] and Florentin [18] applied the generalized inverse in control theory by using its least square properties in the mean square error analysis. Recently, the present authors [3] following on Bott and Duffin [8] have used the g.i. in the analysis of electrical networks, and obtained the explicit solution of a network, d.c. or a.c., in terms of its topological and dynamical characteristics. For extensions of the g.i. and related results to associative rings and semigroups see Drazin [13], Munn and Penrose [31].

**2. Closed operators on a Hilbert space.**<sup>18</sup> For a linear operator  $A$  between two Hilbert spaces  $A: \mathfrak{H}^1 \rightarrow \mathfrak{H}^2$  with  $\overline{D(A)} = \mathfrak{H}^1$  Tseng<sup>19</sup> defined a g.i.  $A^+$  as follows:

*Definition 4.*  $A^+$  is a g.i. of  $A$  if  $\overline{D(A^+)} = \mathfrak{H}^2$  and

$$(25) \quad R(A) \subseteq D(A^+), \quad R(A^+) \subseteq D(A),$$

$$(26) \quad AA^+ = P_{\overline{R(A)}}, \quad A^+A = P_{\overline{R(A^+)}}.$$

Because of the reciprocity in this definition,  $A$  is a g.i. of  $A^+$ , thus  $A = A^{++}$ .

A criterion for the existence and uniqueness of  $A^+$  is:

**THEOREM<sup>20</sup> 6.** *A necessary and sufficient condition for a linear operator  $A$  with  $\overline{D(A)} = \mathfrak{H}^1$  to have a g.i. is that<sup>21</sup>*

$$(27) \quad D(A) = N(A) \oplus \{O(N(A)) \cap D(A)\}.$$

*In this case the operator  $A$  has a unique maximal g.i.  $A_*^+$  (with a maximal*

<sup>18</sup> We refer here only to Tseng's works [45, 46, 47, 48] where the g.i. is treated explicitly. Results related to this subject appear in various other works, e.g. the theory of *pseudo-resolvents* of unbounded linear operators on a  $(B)$ -space to itself as given in Hille and Phillips [23, especially theorem 5.8.4. on p. 186] For another example see Theorem 15 of this paper.

<sup>19</sup> Due to the nature of the publishing journals, proofs are not given by Tseng [45, 46, 47, 48].

<sup>20</sup> This is Theorem A of Tseng [46].

<sup>21</sup> In terms of the orthogonal projection  $P_{\overline{N(A)}}$ , (27) is equivalent to  $P_{\overline{N(A)}}D(A) \subseteq N(A)$ .

domain, i.e. every other g.i. is a restriction of  $A_*^+$ ) with

$$(28) \quad D(A_*^+) = R(A) \oplus O(R(A)), \quad N(A_*^+) = O(R(A))$$

and no other g.i. has a closed nullmanifold.<sup>22</sup>

The following theorem allows a geometrical interpretation of  $A^+$  anticipated by the discussion of the finite dimensional case. Consider the linear equation (here  $A$  is a closed linear operator with dense domain)

$$(29) \quad Ax = a, \quad a \in \mathfrak{R}^2,$$

and its virtual solutions  $x_a$  satisfying

$$(30) \quad \|Ax_a - a\| = \inf_{x \in D(A)} \|Ax - a\|$$

(which need not exist, as is the case for instance if  $\lambda = 0$  is in the continuous spectrum of  $A$  and  $a \notin R(A)$ ). An existence criterion for virtual solutions, and the connection with  $A^+$  are given in:

**THEOREM 7.** *A necessary and sufficient condition that (29) has virtual solutions is that there exists a constant  $G$  such that*

$$(31) \quad |(a, y)|^2 \leq G(y, AA^*y)$$

for every  $y \in D(AA^*) \ominus N(AA^*)$ . In this case the solution<sup>23</sup>

$$(32) \quad x_a^0 = A^+a$$

is of least norm among all the virtual solutions.

Tseng [48] used the Von Neumann canonical decomposition ( $A = WB$  for any closed  $A$  with dense domain, where  $B = \sqrt{A^*A}$  and  $W$  is a partial isometry) in giving a geometrical interpretation of the solution (32) in terms of a "generalized Hesse normal form."

We now summarize some properties of  $A^+$ .

**THEOREM<sup>24</sup> 8.** *For any g.i.  $A^+$  of an operator  $A$ ,*

- (a)  $R(A^+) = A^+R(A) = D(A) \cap O(N(A)) = \overline{P_{\mathfrak{R}(A^+)}}D(A)$ ,  
 $\overline{R(A^+)} = O(N(A))$ ,
- (b)  $N(A) = D(A) \cap O(R(A^+))$ ,  $\overline{N(A)} = O(R(A^+))$ ,
- (c)  $D(A) = N(A) \oplus \{D(A) \cap O(N(A))\}$   
 $= R(A^+) \oplus \{D(A) \cap O(R(A^+))\}$ ,
- (d)  $D(A^*) = R(A^{+*}) \oplus \overline{N(A^+)} = N(A^*) \oplus \{D(A^*) \cap O(N(A^*))\}$ ,
- (e)  $R(A^*) = D(A^{+*}) \cap N(A^{+*})$ ,
- (f)  $N(A^*) = O(R(A)) = N(A_*^+) = \overline{N(A^+)}$ ,

<sup>22</sup> In particular it follows that every closed linear operator with dense domain has a unique closed g.i.

<sup>23</sup> Using (28) it can be shown that (31)  $\Rightarrow a \in D(A_*^+)$ .

<sup>24</sup> This is Theorem C of Tseng [46].

- (g)  $A^{+*} = (A_*^+)^*$ ,
- (h)  $A^{+*}A^* = P_{D(A^*)}$ .

Consider a sequence of equations

$$(33) \quad Ax = a_n, \quad a_n \in \mathfrak{R}^2, \quad n = 1, 2, \dots,$$

having virtual solutions of least norm  $x_{a_n}^0$  for all  $n$ . The following theorem is on the convergence of  $x_{a_n}^0$  as implied by a convergence of  $a_n$  :

**THEOREM<sup>25</sup> 9.** *If*

$$(34) \quad (a_n, b) \rightarrow (a, b)$$

for every  $b \in D(A^*) \ominus N(A^*)$  then:

- (a).  $x_{a_n}^0$  converges weakly if and only if the  $\|x_{a_n}^0\|$  are bounded.
- (b).  $x_{a_n}^0$  converges strongly if and only if

$$(35) \quad \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} (x_{a_m}^0, x_{a_n}^0) = \lim_{n \rightarrow \infty} \|x_{a_n}^0\|^2.$$

In both cases the convergence  $x_{a_n}^0 \rightarrow x_a^0$  is to the virtual solution of least norm of the limit equation  $Ax = a$ . In view of the equivalence between definitions 1 and 3 in the finite dimensional case, and of the discussion below on regular rings, the next theorem is of particular interest:

**THEOREM<sup>26</sup> 10.** *Let  $S: \mathfrak{R}^2 \rightarrow \mathfrak{R}^1$  be a linear operator with  $\overline{D(S)} = \mathfrak{R}^2$  which satisfies*

$$(36) \quad ASA = A, \quad \text{in } D(A),$$

$$(37) \quad P_{O(N(A))}SP_{\overline{R(A)}} = S, \quad \text{in } D(S).$$

Then among the restrictions of  $S$  exist g.i. of  $A$ , one of which,  $S_*$ , has maximal domain and

$$(38) \quad D(S_*) = R(A) \oplus \{D(S) \cap O(R(A))\}.$$

The above selection of Tseng's results [45, 46, 47, 48] does not include his classification of operators and their g.i. relative to the closure of  $D(A)$ ,  $R(A)$ ,  $D(A^*)$  and  $R(A^*)$  given in [46, 47], the necessary and sufficient condition for existence of a bounded g.i. given in [47] and some solvability criteria given in [45].

**3. Regular rings<sup>27</sup>.** Regular rings as introduced by Von Neumann [50] form the algebraic stage in which the analytical show of g.i. takes place. We present here, as a basis for our rather limited applications, some of the

<sup>25</sup> This is section 3° of Tseng [48].

<sup>26</sup> This is Theorem B of Tseng [46].

<sup>27</sup> This section is not essential to the understanding of the remainder of the paper.

results in Von Neumann [50, 51] in the original setting. See also McCoy [28, pp. 143–149].

**THEOREM 11 (and Definition).** *A ring with unit is regular if it satisfies any of the following equivalent statements:*<sup>28</sup>

- (a). *Every principal right ideal  $(a)_r$  has an inverse right ideal.*
- (b). *For every  $a$  there exists an idempotent  $e$  such that  $(a)_r = (e)_r$ .*
- (c). *For every  $a$  there exists an element  $x$  such that  $axa = a$ . (Statements (a) and (b) above have corresponding “left” statements, each of which characterizes regularity.)*

For any ring  $\mathcal{R}$  we denote by  $\mathcal{R}_n$  the ring of all  $n \times n$  matrices with elements from  $\mathcal{R}$ .

**THEOREM<sup>29</sup> 12.** *For every  $n = 1, 2, \dots$ ,  $\mathcal{R}_n$  is regular if and only if  $\mathcal{R}$  is regular.*

The relations between “regularity” and “semi-simplicity”<sup>30</sup> were given by Von Neumann as follows:

**THEOREM 13.** *Let  $\mathcal{R}$  be a ring with unit.*

- (a). *If  $\mathcal{R}$  is regular,  $\mathcal{R}$  is semi-simple.*
- (b). *If  $\mathcal{R}$  is semi-simple and if the chain condition holds for the lattice of all right ideals in  $\mathcal{R}$  then  $\mathcal{R}$  is regular.*

Consider now rings of bounded operators on a complex Hilbert space  $\mathcal{H}$  defined as:  $M$  is a ring if  $A, B \in M \Rightarrow \alpha A, A^*, A + B$  and  $AB \in M$ , and  $M$  is closed in a suitable topology.

**THEOREM 14.** *A necessary and sufficient condition that a ring  $M$  of bounded operators in  $\mathcal{H}$  which contains  $I$  be regular is that  $M$  possess a finite basis<sup>31</sup>.*

We will use now the classification of rings belonging to a factor  $M$  (see definitions 3.1.2 and 4.2.1 in Murray and Von Neumann [32]) according to the range of their relative dimension (see Murray and Von Neumann [32, p. 172], or Naimark [33, p. 469]).

The class  $(I_n)$ ,  $n = 0, 1, 2, \dots$ , is regular by Theorem 14, but this is trivial in view of Theorem 12.<sup>32</sup> For a ring not of a class  $(I_n)$ ,  $n = 0, 1, \dots$  to be regular it is necessary that unbounded operators be adjoined to it. This is possible only for factors of class  $(II_1)$  (see Murray and Von Neumann [32, pp. 122–123]).

<sup>28</sup> We use the terminology and notation of Von Neumann [51].

<sup>29</sup> This theorem suggests the study of g. i. of matrices with elements from  $p$ -rings which are clearly regular. In particular, matrices over Boolean rings are useful in network applications.

<sup>30</sup> Using the terminology of Von Neumann [51, p. 83].

<sup>31</sup> The assumption  $I \in M$  can be avoided, as pointed out in an editorial remark on p. 88 in Von Neumann [51].

<sup>32</sup> See Naimark [33], p. 481 theorem 1.

The following example of such a regular ring is given in Von Neumann [51, p. 89].

Let  $M$  be a factor of class  $(II_1)$  and let  $\mathfrak{u}(M)$  be the set of all linear, closed operators with an everywhere dense domain which "belong" to  $M$ .  $\mathfrak{u}(M)$  is a ring with the closures  $[\alpha A]$ ,  $[A + B]$ ,  $[AB]$  replacing  $\alpha A$ ,  $A + B$ ,  $AB$  in its definition (see Murray and Von Neumann [32, p. 229, Theorem XV]).

**THEOREM 15.** *The ring  $\mathfrak{u}(M)$  is regular.*

The proof of this theorem, given in p. 90 of Von Neumann [51] gives for any  $A \in \mathfrak{u}(M)$  an explicit form for a "g.i.". We reproduce it here:

*Proof.* By theorem 7 in p. 307 of Von Neumann [49],  $A \in \mathfrak{u}(M) \Rightarrow A = WB$  where  $B = \sqrt{A^*A} = W^*A$  is self-adjoint and nonnegative definite, and  $W$  is partially isometric. By lemma 4.4.1 of Murray and Von Neumann [32],  $B, W \in \mathfrak{u}(M)$ .

Define scalar functions

$$(39) \quad f(\lambda) = \begin{cases} 1 & \text{if } \lambda \neq 0, \\ 0 & \text{if } \lambda = 0, \end{cases}$$

$$(40) \quad g(\lambda) = \lambda^+ = \begin{cases} \frac{1}{\lambda} & \text{if } \lambda \neq 0, \\ 0 & \text{if } \lambda = 0. \end{cases}$$

Now  $f(B), g(B) \in \mathfrak{u}(M)$  and:

$$\begin{aligned} f(\lambda)\lambda &= \lambda \Rightarrow [f(B)B] = B, \\ \lambda g(\lambda) &= f(\lambda) \Rightarrow [Bg(B)] = f(B). \end{aligned}$$

Define

$$(41) \quad X = g(B)W^*,$$

then we have

$$\begin{aligned} [AXA] &= [WB \cdot g(B)W^* \cdot A] = [W \cdot Bg(B) \cdot W^*A] \\ &= [W \cdot f(B)B] = [WB] = A. \end{aligned}$$

Thus  $\mathfrak{u}(M)$  is regular.

PART I: PROJECTION PROPERTIES AND EXPANSIONS<sup>33, 34</sup>

**1. A Neumann-Euler expansion for  $A^+$ .** In this section we restrict the discussion, without loss of generality, to square matrices.<sup>35</sup> For a matrix  $A$

<sup>33</sup> We are indebted to Dr. T. N. E. Greville for many helpful suggestions, especially on the interpolation polynomials.

<sup>34</sup> Portions of this part were incorporated in Ben-Israel and Charnes [2] presented at the SIAM meeting, November 1961, Washington, D. C.

<sup>35</sup> See footnote 16.

we denote by  $A^*$ ,  $A^+$  respectively the *conjugate transpose*, and the *Moore-Penrose inverse*. We use here Autonne's canonical form (22) for a square matrix  $A$ :  $A = VDW$  where  $D$  is diagonal,  $V$ ,  $W$  unitary.<sup>36</sup>

The matrix norm  $\|A\|$  most natural for our purposes is

$$(42) \quad \|A\| = \max_{i=1, \dots, n} \{\sqrt{\lambda_i} : \lambda_i \in \sigma(A^*A)\}.$$

This norm is compatible with the Euclidean norm<sup>37</sup> in  $E^n$ .

For applications of a constructive nature (and some theoretical purposes) it is highly desirable to have representations of  $A^+$  in terms of  $A$  and  $A^*$ . Although Moore [29], Bjerhammar [5], Penrose [35] and most recently Greville [20] have given methods for determining the Moore-Penrose inverse, it would be highly desirable to have a series representation analogous to the Neumann expansion for the inverse of a nonsingular matrix<sup>38</sup>.

To date, the only representations of this type available are due to den Broeder and Charnes [11]. These require in an essential manner inverses of matrices like  $A^*A$ . A Neumann type series expansion of  $A^+$  involving only positive powers of  $A^*A$  is given by:

**THEOREM 16.** For any square matrix  $A \neq 0$  and a real number  $\alpha$  with

$$(43) \quad 0 < \alpha < \min_{\substack{d_{ii} \neq 0 \\ i=1, \dots, n}} \frac{2}{|d_{ii}|^2}.$$

where  $d_{ii}$  are the (diagonal) elements of  $D$  in the Autonne's representation (22) of  $A$ , the series

$$\alpha \sum_{k=0}^m (I - \alpha A^*A)^k A^*$$

converges and

$$(44) \quad \alpha \sum_{k=0}^{\infty} (I - \alpha A^*A)^k A^* = A^+$$

*Proof.* Penrose used (22) to write

$$(45) \quad A^+ = W^*D^+V^*,$$

where<sup>39</sup>

$$(46) \quad D^+ = (d_{ii}^+).$$

<sup>36</sup> Extended to rectangular matrices in Hestenes [21, p. 74, theorem 10.3].

<sup>37</sup> For a discussion of matrix norms see Faddeeva [16, pp. 55-60].

<sup>38</sup> E.g., Altman [1, p. 56, theorem 2].

<sup>39</sup> For a scalar  $\lambda$ ,  $\lambda^+ = \begin{cases} \frac{1}{\lambda} & \text{if } \lambda \neq 0, \\ 0 & \text{if } \lambda = 0. \end{cases}$

Using (22) we obtain

$$(47) \quad (I - \alpha A^*A)^k = W^*(I - \alpha D^*D)^k W, \quad k = 0, 1, 2, \dots,$$

$$(48) \quad (I - \alpha A^*A)^k A^* = W^*(I - \alpha D^*D)^k D^*V^*, \quad k = 0, 1, 2, \dots.$$

Using (43) we get

$$(49) \quad \sum_{k=0}^{\infty} (1 - \alpha |d_{ii}|^2)^k d_{ii}^* = \alpha^{-1} d_{ii}^+, \quad i = 1, \dots, n,$$

so that

$$(50) \quad \alpha \sum_{k=0}^{\infty} (I - \alpha A^*A)^k A^* = W^*D^+V^*.$$

*Remarks.*

(i). For nonsingular  $A$ , (44) is the Neumann expansion of  $A^{-1}$  written as  $A^{-1} = (A^*A)^{-1}A^*$ , see for example Altman [1, p. 56].

(ii). Comparing (44) and the den Broeder-Charnes equation (23) and using the uniqueness of  $A^+$  we obtain the interesting Moebius type matrix identity:

COROLLARY 2. For any square matrix  $A$  and a real number  $\alpha$  as in (43),

$$(51) \quad \sum_{k=1}^{\infty} A^*(I + AA^*)^{-k} = \alpha \sum_{k=0}^{\infty} (I - \alpha A^*A)^k A^*.$$

The Neumann expansion (44) for  $A^+$  can be replaced by a more rapidly convergent scheme, making use of an identity due to Euler,

$$(52) \quad \frac{1}{1-x} = (1+x)(1+x^2)(1+x^4)(1+x^8) \dots, \quad \text{for } |x| < 1,$$

a finite abridgement of which is <sup>40</sup>

$$(53) \quad (1+x) \prod_{k=1}^{n-1} (1+x^{2^k}) = \sum_{k=0}^{2^n-1} x^k, \quad \text{for } |x| < 1.$$

A Neumann-Euler expansion for  $A^+$  is given in:

THEOREM 17. For any square matrix  $A \neq 0$  and a real number  $\alpha$  as in (43), let

$$(54) \quad A_n^+ = \alpha \{I + (I - \alpha A^*A)\} \prod_{k=1}^{n-1} \{I + (I - \alpha A^*A)^{2^k}\} A^*.$$

Then

$$(55) \quad \|A^+ - A_n^+\| \leq \frac{(1 - \alpha d_{11}^2)^{2^n}}{|d_{11}|}.$$

<sup>40</sup> See also Lonseth [26].

where

$$(56) \quad |d_{11}| = \max_{i=1, \dots, n} |d_{ii}|,$$

and  $\{d_{ii}\}_{i=1}^n$  are the elements of  $D$  in (22).

*Proof.* Using (44), (52) and (53) we get

$$(57) \quad A^+ - A_n^+ = \alpha \sum_{k=2^n}^{\infty} (I - \alpha A^* A)^k A^*,$$

or using (22),

$$(58) \quad A^+ - A_n^+ = \alpha W^* \sum_{k=2^n}^{\infty} (I - \alpha D^* D)^k D^* V^*.$$

Since the norm (42) is multiplicative<sup>41</sup>, we have

$$(59) \quad \begin{aligned} \|A - A_n^+\| &\leq \alpha \|W^*\| \sum_{k=2^n}^{\infty} \|(I - \alpha D^* D)^k D^*\| \|V^*\| \\ &= \alpha \sum_{k=2^n}^{\infty} \|(I - \alpha D^* D)^k D^*\| \end{aligned}$$

from which (55) follows immediately.

Similarly, other rapidly convergent iterative schemes for  $A^+$  may be constructed by using the generalizations of (52) suggested by Lonseth in [26].

**2. Projections associated with the Moore-Penrose inverse.** Penrose in [34] showed that  $AA^+$ ,  $A^+A$  are hermitian idempotents, thus projections. In making these statements more precise we establish the following theorem which is the basis of Moore's definition of  $A^+$  and of the "geometrical" definition given by Hestenes in [24, pp. 84-85].

**THEOREM 18.** For any  $m \times n$  matrix  $A$ , (a)  $AA^+$  is a projection on  $R(A)$  along  $N(A^*)$ , and (b)  $A^+A$  is a projection on  $R(A^*)$  along  $N(A)$ .

*Proof.* (a). By (7) and (9)  $AA^+$  is hermitian idempotent, and it remains to characterize the range of the projection  $AA^+$  and its null space<sup>42</sup>.

(i). Because of (7),  $R(A) \subseteq R(AA^+)$ . Conversely if

$$(60) \quad AA^+x = x, \quad \text{for any } x \in E^m,$$

then clearly  $x \in R(A)$ .

Thus we have established  $R(AA^+) \subseteq R(A)$ , and hence<sup>43</sup>

$$(61) \quad R(A) = R(AA^+).$$

<sup>41</sup> i.e.,  $\|AB\| \leq \|A\| \|B\|$ .

<sup>42</sup> The range of a projection  $P$  in  $E^n$  is  $R(P) = PE^n = \{x: x \in E^n, Px = x\}$ . The null space of  $P$  is  $N(P) = R(I - P)$ . We say that  $P$  is a projection on  $R(P)$  along  $N(P)$ .

<sup>43</sup> This is a corollary of theorem 2 in Penrose [34].

(ii). It remains to show that

$$(62) \quad N(AA^+) = N(A^*),$$

or equivalently that

$$(63) \quad R(I - AA^+) = N(A^*).$$

Now for any  $x \in R(I - AA^+)$  we have

$$(64) \quad x = (I - AA^+)x$$

and<sup>44</sup>

$$(65) \quad \begin{aligned} A^*x &= (A^* - A^*AA^+)x = (A^* - A^*(AA^+)^*)x \\ &= (A^* - A^*A^{**}A^*)x = 0. \end{aligned}$$

Hence  $R(I - AA^+) \subseteq N(A^*)$ .

Conversely, for any  $x \in N(A^*)$  we have

$$(66) \quad 0 = A^{**}(A^*x) = (AA^+)^*x = AA^+x,$$

so that  $N(A^*) \subseteq N(AA^+)$  and the proof of (62), and of part (a) is completed.

(b). The proof of (b) follows from that of (a) by noting that

$$A^+A = (A^+A)^* = A^*A^{**}.$$

A corollary to Theorem 18 sheds light on the structural relations between  $A^*$  and  $A^+$  as given originally by Moore in [29, 30].

COROLLARY 3. (a). *The null spaces of  $A^*$ ,  $A^+$  are identical, i.e.,*

$$(67) \quad N(A^*) = N(A^+).$$

(b). *The ranges of  $A^{**}$ ,  $A$  are identical, i.e.,*

$$(68) \quad R(A^{**}) = R(A).$$

*Proof.* (a). That  $N(A^+) \subseteq N(A^*)$  follows immediately from (62). The converse follows from (62) by using (9). (b). This is contained in the statement of the next corollary, using (67).

In the proof of Theorem 18 we used only the defining relations (7), (8), (9) and (10), the existence of  $A^+$  and the characterization of hermitian idempotents as orthogonal projections. As a corollary to Theorem 18 we obtain a decomposition theorem for  $E^n$  in terms of suitable range and null spaces, which is the basis for Fredholm's alternative for matrix equations and linear integral equations with degenerate kernel.

<sup>44</sup> In (65) we made use of (9) and Lemma 1 (b), although we could directly use equation (10) in Penrose [34].

COROLLARY 4. For any  $m \times n$  matrix  $A$

- (a)  $E^m = R(A) \oplus N(A^*),$
- (b)  $E^n = R(A^*) \oplus N(A).$

*Proof.* Parts (a) and (b) are restatements of parts (a) and (b) respectively of Theorem 18.

Let now  $f, f', g, g'$  denote vectors in  $E^n$  and  $K$  a square matrix.

COROLLARY 5. (*Fredholm*). Consider

- (69)  $f - Kf = g$
- (70)  $f - Kf = 0$
- (71)  $f' - K^*f' = g'$
- (72)  $f' - K^*f' = 0$

Then (a). The numbers of linearly independent nontrivial solutions of (70) and (72) are equal, (b). (i), [(ii)]. (69) [(71)] is<sup>45</sup> solvable if and only if  $g[g']$  is orthogonal to every solution of (72) [(70)] in which case the solution is determined up to a linear combination of the solutions of (70) [(72)], and (c). (i), [(ii)]. (69) [(71)] is uniquely solvable if (72) [(70)] has no trivial solutions.

*Proof.* Write  $I - K = T$ . (a). To prove (a) is equivalent to showing that  $T$  and  $T^*$  have the same rank, which is a well known fact. (If we choose to prove (a) by using Theorem 18, it is equivalent to showing that  $TT^+$  and  $T^+T$  have the same rank, which is Lemma 1 (h)).

(b), (i). By Theorem 18 (a),  $g \in R(T)$  if and only if  $(I - TT^+)g = 0$ , i.e.,  $g$  is orthogonal to  $N(T^*)$ . Then, by (6),  $f = T^+g + (I - T^+T)z$  where  $z \in E^n$  is arbitrary. By Theorem 18 (b) this implies that  $f$  is determined only up to linear combinations of elements of  $N(T)$ .

(b), (ii). This is proved similarly.

(c). This follows immediately as a corollary of (b).

In proving Theorem 3, Penrose showed that the spectral representation of  $AA^*$ ,

$$(73) \quad AA^* = \sum_{\lambda \geq 0} \lambda E_\lambda$$

has as its projection family

$$(74) \quad E_\lambda = I - (AA^* - \lambda I)^+(AA^* - \lambda I), \quad \lambda \geq 0.$$

From this, by using Theorem 18, we conclude a characterization of  $R(A)$  in terms of the eigenvectors of  $AA^*$ .

<sup>45</sup> Parts (b) and (c) contain two statements each, denoted by (i), [(ii)], with square brackets in the statements applying to (ii).

**THEOREM 19.** For any  $m \times n$  matrix  $A$

$$(75) \quad R(A) = \bigoplus_{\lambda > 0} N\{AA^* - \lambda I\},$$

i.e. the range of  $A$  is the direct sum of the eigenmanifolds of  $AA^*$  corresponding to its nonzero eigenvalues.<sup>46</sup>

*Proof.* It follows from lemma 1.5 and equation (10) of Penrose [34] that

$$(76) \quad AA^+ = (AA^*)^+AA^*.$$

Using (76), (73) and (74) we get

$$(77) \quad AA^+ = \sum_{\lambda > 0} E_\lambda,$$

and (75) results from the following facts:

- (i).  $AA^+$  is a projection on  $R(A)$ .
- (ii).  $E_\lambda$  is a projection on  $N\{AA^* - \lambda I\}$ .
- (iii).  $N\{AA^* - \lambda I\}, N\{AA^* - \mu I\}$  are orthogonal subspaces for  $\lambda \neq \mu$ .

As an illustration of the above methods we consider now the vector space  $E^{m \times n}$  of all  $m \times n$  complex matrices over the complex field, with two associated rings of matrix operators,  $\Theta_l$  and  $\Theta_r$ , operating on every  $X \in E^{m \times n}$  as pre- and post-multipliers respectively. Taking  $\Theta_l = E^{m \times m}$  and  $\Theta_r = E^{n \times n}$ , we have  $\Theta_l E^{m \times n} \Theta_r = E^{m \times n}$ . We denote then the general dyadic operator<sup>47</sup> by  $(A \bullet \bullet B)$  where  $A \in E^{m \times m}, B \in E^{n \times n}$ , meaning that

$$(78) \quad (A \bullet \bullet B)X = AXB, \quad \text{for } X \in E^{m \times n}.$$

The reason for looking on  $E^{m \times n}$  this way rather than as an Euclidean space  $E^{mn}$  is that even without defining inner products (thus orthogonality) in  $E^{m \times n}$  we can state explicitly a decomposition theorem for  $E^{m \times n}$  relative to its multiplier rings  $E^{m \times m}$  and  $E^{n \times n}$  which is an extension of Corollary 4. (Indeed Corollary 4 is a special case of Theorem 20 below, by setting  $n = 1, B = I$  and using Corollary 3 (a).)

By a subspace in  $E^{m \times n}$  we mean a subset of  $E^{m \times n}$  which is closed under complex linear combinations of its elements. Clearly for any  $A \in E^{m \times m}, B \in E^{n \times n}$  the range and nullspace of  $(A \bullet \bullet B)$  defined respectively as

$$(79) \quad R(A \bullet \bullet B) = \{X: X \in E^{m \times n}, X = AUB \text{ for some } U \in E^{m \times n}\},$$

$$(80) \quad N(A \bullet \bullet B) = \{X: X \in E^{m \times n}, AXB = 0\},$$

<sup>46</sup> Or in a more familiar form,  $R(A) = R(AA^*)$ .

<sup>47</sup> Our usage of the word "dyadic" here is not in accordance with the established terminology, e.g. Hitchcock [55], where the equation  $A_1XB_1 + \dots + A_nXB_n = C$  is solved by using Gibbs' dyadics. Compare with Penrose [34] where the same equation is reduced to an ordinary vector equation in a space of higher dimension and solved by using (6).

are subspaces. Finally it remains to explain what we mean by a *direct sum* of subspaces in  $E^{m \times n}$ . We write

$$(81) \quad L = M \oplus N$$

and say that the subspace  $L$  is a direct sum of the subspaces  $M$  and  $N$  if every  $X \in L$  can be uniquely represented as

$$(82) \quad X = Y + Z, \quad Y \in M, \quad Z \in N.$$

Understanding the product of two dyadic operators  $(A \cdot \cdot B)$  and  $(C \cdot \cdot D)$ ,  $A, C \in E^{m \times m}$ ,  $B, D \in E^{n \times n}$  in that order, as

$$(83) \quad (A \cdot \cdot B)(C \cdot \cdot D) = (AC \cdot \cdot DB),$$

we have the following lemma.

LEMMA 2. For any  $A \in E^{m \times m}$ ,  $B \in E^{n \times n}$ , the dyadic operators

$$(a) \quad (A^+ A \cdot \cdot B B^+),$$

and

$$(b) \quad (A A^+ \cdot \cdot B^+ B)$$

are idempotent.

*Proof.* Obvious by (7) or (8).

In analogy to Theorem 18 it is thus true that the operators

$$(A^+ A \cdot \cdot B B^+), \quad (I \cdot \cdot I) - (A^+ A \cdot \cdot B B^+),$$

$$(A A^+ \cdot \cdot B^+ B), \quad (I \cdot \cdot I) - (A A^+ \cdot \cdot B^+ B),$$

are *parallel projections* (another word for idempotents). The ranges of the pair

$$(A A^+ \cdot \cdot B^+ B), \quad (I \cdot \cdot I) - (A A^+ \cdot \cdot B^+ B)$$

are characterized in the following decomposition theorem which is a consequence of Theorem 1 and Lemma 2.

THEOREM 20. For any  $A \in E^{m \times m}$ ,  $B \in E^{n \times n}$ ,

$$(84) \quad E^{m \times n} = R(A \cdot \cdot B) \oplus N(A^+ \cdot \cdot B^+).$$

*Proof.* Write any  $X \in E^{m \times n}$  as (82) with

$$(85) \quad Y = A A^+ X B^+ B,$$

$$(86) \quad Z = X - A A^+ X B^+ B.$$

(i). By Theorem 1 and Lemma 2(b),

$$Y \in R(A \cdot \cdot B).$$

(ii). That  $Z \in N(A^+ \cdot \cdot B^+)$  follows from Lemma 2(b).

(iii). To complete the proof one has to show that  $X = 0$  is the only

matrix  $\in E^{m \times n}$  common to  $R(A \bullet \bullet B)$  and to  $N(A^+ \bullet \bullet B^+)$ . Suppose this is not so, then there exists  $X \neq 0 \ni X \in R(A \bullet \bullet B) \cap N(A^+ \bullet \bullet B^+)$ . By Theorem 1,  $X = AA^+XB^+B$  so that  $X \in N(A^+ \bullet \bullet B^+) \Rightarrow X = 0$ , a contradiction.<sup>48</sup>

*Remarks.*

(i). These results suggest applications to integral equations of the type

$$(87) \quad \int_a^b \int_a^b A(r, s)X(s, t)B(t, u) = C(r, u),$$

where  $A(r, s)$  and  $B(t, u)$  are degenerate kernels.

(ii). Defining in  $E^{m \times n}$  the norm  $\|X\|_1 = \sqrt{\text{trace}(X^*X)}$  is equivalent to defining in  $E^{m \times n}$  the inner product<sup>49</sup>  $(X, Y)_1 = \sum_{i=1}^m \sum_{j=1}^n x_{ij}y_{ij}$ . As  $\|X\|_1$  is simply the Euclidean norm when  $X$  is considered as a vector in  $E^{mn}$  rather than in  $E^{m \times n}$ , we expect a "least square" interpretation of the decomposition given in Theorem 20, by using the norm  $\|\cdot\|_1$ . In fact, this is the theorem in Penrose [35] which we restate below, using our terminology: For any given pair  $A \in E^{m \times m}$ ,  $B \in E^{n \times n}$  we consider for every  $X \in E^{m \times n}$  the (always nonempty) set of all  $(A, B)$ -representations  $\{X_1, X_2\}$  where  $X_1, X_2 \in E^{m \times n}$  and  $X = AX_1B + X_2$ . The  $(A, B)$ -representation  $\{A^+XB^+, X - AA^+XB^+B\}$  given in Theorem 20, has the following "minimal" character. For any other representation  $\{X_1, X_2\}$ , either

$$\|X - AA^+XB^+B\|_1 < \|X_2\|_1,$$

or

$$\|X - AA^+XB^+B\|_1 = \|X_2\|_1 \quad \text{and} \quad \|A^+XB^+\|_1 \leq \|X_1\|_1.$$

Thus  $A^+XB^+$ , the *least square solution*<sup>50</sup> of  $AUB = X$ , is shown to correspond to the projection  $AA^+XB^+B$  of  $X$  on  $R(A \bullet \bullet B)$ .

**3. An interpolation polynomial for the Moore-Penrose inverse.** In this section we express  $A^+$  as a Lagrange-Sylvester interpolation polynomial in powers of  $A, A^*$ . For any complex square matrix<sup>51</sup>  $A$  let  $\sigma(A)$  denote the *spectrum* of  $A$ , and  $\psi(A)$  its *minimal polynomial* written as

$$(88) \quad \psi(\lambda) = \prod_{\theta \in \sigma(A)} (\lambda - \theta)^{\nu(\theta)},$$

where the root  $\theta \in \sigma(A)$  is *simple* if  $\nu(\theta) = 1$  and *multiple* otherwise.

For any scalar function  $f(\lambda)$  which is analytic at the multiple roots of

<sup>48</sup> Without defining *inner products* it is not possible to talk about the *orthogonality* of the ranges  $R(P), R(I - P)$  of two complementary parallel projections. However,  $R(P) \oplus R(I - P) = E^{m \times n}$ , in the sense of (81) and (82).

<sup>49</sup> Indeed  $\|X\|_1 = (X, X)_1^{1/2}$ . This norm is due to Wedderburn [56].

<sup>50</sup> Also called "best approximate solution" by Penrose [35], "extremal virtual solution" by Tseng [45,48].

<sup>51</sup> See footnote 16.

$\psi(\lambda)$  and defined at the simple roots of  $\psi(\lambda)$  it is possible to construct a *matrix function*  $f(A)$  which satisfies the first four Fantappie requirements:<sup>52</sup>

- (I).  $f(\lambda) = k \Rightarrow f(A) = kI,$
- (II).  $f(\lambda) = \lambda \Rightarrow f(A) = A,$
- (III).  $f(\lambda) = g(\lambda) + h(\lambda) \Rightarrow f(A) = g(A) + h(A),$
- (IV).  $f(\lambda) = g(\lambda)h(\lambda) \Rightarrow f(A) = g(A)h(A).$

We intend to construct  $A^+$  as the matrix function  $f(A)$  corresponding to the scalar function  $f(\lambda) = \lambda^+$  and consider only the case where  $\lambda = 0 \in \sigma(A)$  as otherwise  $A$  is nonsingular (thus uninteresting for us). Even if  $\lambda = 0$  is a simple root, the attempts to construct  $A^+$  this way lead only to the satisfaction of (7) and (8) but generally not of (9) or<sup>53</sup> (10). This difficulty clearly does not exist if  $A$  is Hermitian. For a Hermitian matrix represented as

$$(89) \quad H = \sum \lambda E_\lambda,$$

Tseng in [44] showed that the g.i. is<sup>54</sup>

$$(90) \quad H^+ = \sum \lambda^+ E_\lambda.$$

We use therefore equation (10) in Penrose [34],

$$(91) \quad A^+ = (A^*A)^+A^*,$$

to construct  $A^+$  by using the Lagrange-Sylvester interpolation polynomial to give explicit expressions for the projections  $E_\lambda$  (see (89)) associated with  $(A^*A)$ , as all the roots in  $\sigma(A^*A)$  are simple.<sup>55</sup>

$$(92) \quad A^*A = \sum_{\lambda \in \sigma(A^*A)} \lambda \left( \frac{\prod_{\substack{\lambda \neq \theta \in \sigma(A^*A)}} (A^*A - \theta I)}{\prod_{\lambda \neq \theta \in \sigma(A^*A)} (\lambda - \theta)} \right),$$

so that by (89), (90) and (91),

$$(93) \quad A^+ = \sum_{\lambda \in \sigma(A^*A)} \lambda^+ \left( \frac{\prod_{\lambda \neq \theta \in \sigma(A^*A)} (A^*A - \theta I)}{\prod_{\lambda \neq \theta \in \sigma(A^*A)} (\lambda - \theta)} \right) A^*.$$

We call (93) the *Lagrange-Sylvester interpolation polynomial*<sup>56, 57</sup> for  $A^+$ .

<sup>52</sup> See Fantappie [17], MacDuffee [27, pp. 99-102] and Schwerdtfeger [43, p. 30]. For a review of this subject and a comparison of the various definitions see Rinehart [40].

<sup>53</sup> Since the function  $f(\lambda) = \bar{\lambda}$  is not analytic, the above correspondence  $f(\lambda) \leftrightarrow f(A)$  does not include  $\bar{\lambda} \leftrightarrow A^*$ .

<sup>54</sup> See the generalization by Penrose in theorem 3, equations (17), (20).

<sup>55</sup> E. g., MacDuffee [27, p. 99 bottom].

Examples. (1). Let

$$A = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}.$$

Then

$$A^*A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},$$

and

$$\psi(A^*A) = (A^*A)^2 - 2(A^*A)$$

is the minimal polynomial. Writing

$$\psi(\lambda) = \lambda(\lambda - 2),$$

we have

$$(A^*A)^+ = \frac{1}{2} \begin{pmatrix} A^*A \\ -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{pmatrix},$$

and

$$A^+ = (A^*A)^+A^* = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix} \frac{1}{2} A^*.$$

(2). Let

$$A = \begin{pmatrix} 0 & e^{i\theta} & 0 \\ e^{i\varphi} & 0 & e^{i\varphi} \\ 0 & e^{i\theta} & 0 \end{pmatrix}.$$

Then

$$A^* = \begin{pmatrix} 0 & e^{-i\varphi} & 0 \\ e^{-i\theta} & 0 & e^{-i\theta} \\ 0 & e^{-i\varphi} & 0 \end{pmatrix}, \quad A^*A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

The minimal polynomial of  $A^*A$  is  $\psi(\lambda) = \lambda(\lambda - 2)$ .

<sup>56</sup> An improvement over (93) is the following result, communicated to us by a referee: Let  $q(\lambda)$  be chosen so that

$$\psi(\lambda) = c(\lambda^k - \lambda^{k+1}q(\lambda)), \quad c \neq 0.$$

Then

$$A^+ = A^*q(AA^*) = q(A^*A)A^*.$$

<sup>57</sup> Equation (93) is not a practical way for computing  $A^+$ , for it is very sensitive to errors in the computed values of  $\sigma(A^*A)$ .

$$(A^*A)^+ = \frac{1}{2} \left( \frac{A^*A}{2} \right) = \frac{1}{4} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$$A^+ = \frac{1}{4} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & e^{-i\varphi} & 0 \\ e^{-i\theta} & 0 & e^{-i\theta} \\ 0 & e^{-i\varphi} & 0 \end{pmatrix} \\ = \frac{1}{2} \begin{pmatrix} 0 & e^{-i\varphi} & 0 \\ e^{-i\theta} & 0 & e^{-i\theta} \\ 0 & e^{-i\varphi} & 0 \end{pmatrix}.$$

(3). Let

$$A = \begin{pmatrix} 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix},$$

$$A^*A = \begin{pmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \end{pmatrix}.$$

The minimal polynomial of  $(A^*A)$  is

$$\psi(\lambda) = \lambda(\lambda - 2)(\lambda - 4).$$

Therefore,

$$(A^*A)^+ = \frac{1}{2} \left( \frac{A^*A(A^*A - 4I)}{2(2 - 4)} \right) + \frac{1}{4} \left( \frac{A^*A(A^*A - 2I)}{4(4 - 2)} \right) \\ = \frac{14}{32} (A^*A) - \frac{3}{32} (A^*A)^2 = \frac{1}{16} \begin{pmatrix} 5 & -1 & -3 & -1 \\ -1 & 5 & -1 & -3 \\ -3 & -1 & 5 & -1 \\ -1 & -3 & -1 & 5 \end{pmatrix}.$$

Therefore,

$$A^+ = \frac{1}{16} \begin{pmatrix} 5 & -1 & -3 & -1 \\ -1 & 5 & -1 & -3 \\ -3 & -1 & 5 & -1 \\ -1 & -3 & -1 & 5 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 1 \end{pmatrix} \\ = \frac{1}{8} \begin{pmatrix} 3 & -3 & -1 & 1 \\ 1 & 3 & -3 & -1 \\ -1 & 1 & 3 & -3 \\ -3 & -1 & 1 & 3 \end{pmatrix}.$$

PART II: ON THE MOORE-TSENG INVERSE OF A BOUNDED SYMMETRIC OPERATOR ON A HILBERT SPACE INTO ITSELF<sup>58</sup>

Following a brief discussion of the general case, the g.i. of a bounded symmetric linear operator on a Hilbert space into itself is explicitly given and some of its properties are established. These results can be extended to closed unbounded operators whose domain is everywhere dense.

**1. The g.i. of a linear operator between Hilbert spaces<sup>59</sup>.** Let  $A: \mathcal{H}^1 \rightarrow \mathcal{H}^2$  be a *linear transformation* from a Hilbert space  $\mathcal{H}^1$  to a Hilbert space  $\mathcal{H}^2$ . The *Cartesian product*

$$(94) \quad \mathcal{H} = \mathcal{H}^1 \times \mathcal{H}^2$$

is the class of all pairs  $\{x, y\}$  with  $x \in \mathcal{H}^1, y \in \mathcal{H}^2$ .  $\mathcal{H}$  is a Hilbert space with the induced inner product

$$(95) \quad (\{x_1, y_1\}, \{x_2, y_2\}) = (x_1, x_2) + (y_1, y_2),$$

for  $x_1, x_2 \in \mathcal{H}^1, y_1, y_2 \in \mathcal{H}^2$ . In  $\mathcal{H}$  we identify  $\mathcal{H}^1$  with the subspace  $\mathcal{H}^{(1)}$  of all points  $\{x, 0\}$ , and  $\mathcal{H}^2$  with the subspace  $\mathcal{H}^{(2)}$  of all points  $\{0, y\}$ .

The *graph* of  $A: \mathcal{H}^1 \rightarrow \mathcal{H}^2$  in  $\mathcal{H}$ , denoted by  $G(A)$ , is the subspace  $\{x, Ax\}$  with  $x \in D(A)$ . Clearly  $G(A) \cap \mathcal{H}^{(2)} = 0$ . Conversely, a subspace  $\mathcal{G}$  of  $\mathcal{H}$  defines a linear transformation  $A: \mathcal{H}^1 \rightarrow \mathcal{H}^2$  if the points of  $\mathcal{G}$  have a distinguishing first element, that is, if

$$(96) \quad \mathcal{G} \cap \mathcal{H}^{(2)} = 0.$$

Then  $\mathcal{G} = G(A)$ , and the domain and range of  $A$  are respectively<sup>60</sup>

$$(97) \quad D(A) = J_1^{-1}P_{\mathcal{H}^{(1)}}\mathcal{G},$$

$$(98) \quad R(A) = J_2^{-1}P_{\mathcal{H}^{(2)}}\mathcal{G},$$

where  $J_1, J_2$  denote the isomorphisms respectively mapping  $\mathcal{H}^1$  on  $\mathcal{H}^{(1)}$ , and  $\mathcal{H}^2$  on  $\mathcal{H}^{(2)}$ .

The linear transformation  $A$  is *closed* if  $\mathcal{G}$  is closed in  $\mathcal{H}$ . Moreover,  $A$  has a *dense domain* if  $\overline{D(A)} = \mathcal{H}^1$  or equivalently if  $O(\mathcal{G}) \cap \mathcal{H}^{(1)} = 0$  where  $O(\mathcal{G})$  is the orthogonal complement of  $\mathcal{G}$  in  $\mathcal{H}$ .

Suppose that  $\mathcal{G}$  is closed. Then the classes

$$(99) \quad \mathcal{G}_0 = \mathcal{G} \cap \mathcal{H}^{(1)} = J_1N(A)$$

<sup>58</sup> This title is suggested by the history of the subject. However Tseng's articles [45, 46, 47, 48] became known to us only after this part was written. Our approach is different from Tseng's as spectral theory is used here to give an explicit construction of  $A^+$ , for this more restricted case.

<sup>59</sup> The material in this section was communicated to us by one of the referees, Magnus Hestenes, who has identified himself at our insistence.

<sup>60</sup> E.g., Von Neumann [52, p. 59].

and

$$(100) \quad \mathfrak{B} = \mathfrak{A} \cap O(\mathfrak{A}_0)$$

are mutually orthogonal and  $\mathfrak{A}$  is the direct sum

$$(101) \quad \mathfrak{A} = \mathfrak{A}_0 \oplus \mathfrak{B}.$$

Set

$$(102) \quad \mathfrak{A}_0^+ = O(\mathfrak{A}) \cap \mathfrak{H}^{(2)},$$

$$(103) \quad \mathfrak{A}^+ = \mathfrak{A}_0^+ \oplus \mathfrak{B}.$$

The class  $\mathfrak{A}^+$  satisfies

$$(104) \quad \mathfrak{A}^+ \cap \mathfrak{H}^{(1)} = 0,$$

by construction.

The elements of  $\mathfrak{A}^+$  accordingly have a distinguishing second element and thus define a linear transformation from  $\mathfrak{H}^2$  to  $\mathfrak{H}^1$ .

*Definition 5.* The *generalized inverse*  $A^+ : \mathfrak{H}^2 \rightarrow \mathfrak{H}^1$  of the closed linear operator  $A : \mathfrak{H}^1 \rightarrow \mathfrak{H}^2$  is the linear transformation whose graph is  $\mathfrak{A}^+$  in (103). That is

$$(105) \quad \mathfrak{A}^+ = G(A^+).$$

This definition makes sense even if the domain of  $A$  is not dense. We will, however, restrict the discussion to the case in which  $D(A) = \mathfrak{H}^1$ . In this event

$$(106) \quad O(\mathfrak{A}) \cap \mathfrak{H}^{(1)} = 0,$$

so that  $O(\mathfrak{A})$  defines a linear transformation from  $\mathfrak{H}^2$  to  $\mathfrak{H}^1$ . To relate this transformation to  $A$  we note that its graph  $O(\mathfrak{A})$  is, by definition, the orthogonal complement of  $\mathfrak{A} = G(A)$ . Therefore<sup>61</sup>

$$(107) \quad O(\mathfrak{A}) = G(-A^*),$$

where  $A^*$  is the adjoint of  $A$ .

Let the subspace  $\mathfrak{C}$  be defined by

$$(108) \quad \mathfrak{C} = O(\mathfrak{A}) \ominus \mathfrak{A}_0^+,$$

where  $\mathfrak{A}_0^+$  is as in<sup>62</sup> (102). Using definition 5 to define the g.i. of  $-A^*$ , we verify that its graph is

$$(109) \quad (O(\mathfrak{A}))^+ = \mathfrak{A}_0 \oplus \mathfrak{C}.$$

<sup>61</sup> E.g., Von Neumann [52, definition 13.13, p. 62].

<sup>62</sup> Equivalent to  $O(\mathfrak{A}) = \mathfrak{C} \oplus \mathfrak{A}_0^+$ .

**Definition 6.** For a closed linear operator  $A: \mathfrak{H}^1 \rightarrow \mathfrak{H}^2$  whose domain is dense, the *generalized inverse*,  $(-A^*)^+$ , of  $-A^*: \mathfrak{H}^2 \rightarrow \mathfrak{H}^1$  is the linear transformation from  $\mathfrak{H}^1$  to  $\mathfrak{H}^2$  whose graph is  $(O(\mathfrak{G}))^+$  in (109).

We easily verify that<sup>63</sup>  $(-A^*)^+ = -A^{+*}$ . Indeed, most of the properties of the generalized inverse follow easily from definitions 5 and 6.  $A^+$ , defined by definition 5, coincides with Tseng's *maximal g.i.*  $A_*^+$ , see Theorem 6. In what follows, however, the discussion will be restricted to the case  $\mathfrak{H}^1 = \mathfrak{H}^2$  where  $A$  is bounded and symmetric. In this case explicit formulae for the g.i. will be given.

**2. The g.i. of a bounded symmetric operator on a Hilbert space into itself.** In this section we make essential use of a theorem by Dunford ([14, Theorem 3.6]) which is reproduced below, following some notations and definitions.

Let  $T$  denote a bounded linear operator on a complex Banach space  $\mathfrak{X}$ , and let  $R_\lambda(T)$  denote the resolvent of  $T$ ,  $\rho(T)$  denote the resolvent set of  $T$ ,  $\sigma(T)$  denote the spectrum of  $T$ .

**Definition 7.** A  $T$ -admissible domain is an open set  $D$  in the complex plane having the following properties: (i),  $D$  is a finite union of connected open sets whose closures are pairwise disjoint, and (ii), the boundary  $C$  of  $D$  consists of a finite number of disjoint rectifiable Jordan curves lying in  $\rho(T)$ . The fact that  $D$  is  $T$ -admissible will be denoted by  $D = D(T)$ .

**Definition 8.**  $\mathfrak{F}(T)$  is the class of all complex functions which are regular and single-valued on the closure of some  $D = D(T)$  which contains  $\sigma(T)$ .

**Definition 9.** A spectral set of  $T$  is any subset  $\sigma$  of  $\sigma(T)$  which is both open and closed in  $\sigma(T)$ .

The following theorem gives necessary and sufficient conditions for a sequence of polynomials,  $f_n(T)$ , to converge uniformly to a projection  $P$ .

**THEOREM 21.** (Dunford) Let

$$F(\lambda) = \prod_{i=1}^k (\lambda - \lambda_i)^{\nu_i}$$

be a polynomial whose distinct roots are  $\lambda_1, \dots, \lambda_k$ . Let  $f_n \in \mathfrak{F}(T)$  satisfy<sup>64</sup>

- (i)  $f_n(\lambda_i) \rightarrow 1, \quad f_n^{(j)}(\lambda_i) \rightarrow 0, \quad \begin{cases} i = 1, \dots, k, \\ j = 1, \dots, \nu_i - 1, \end{cases}$
- (ii)  $F(T)f_n(T) \rightarrow 0.$

Then the following statements are equivalent.

- (a).  $f_n(T) \rightarrow P, P^2 = P, P\mathfrak{X} = N\{F(T)\}.$
- (b). Each  $\lambda_i, i = 1, \dots, k$ , is either in  $\rho(T)$  or else a pole of  $R_\lambda(T)$ .

<sup>63</sup> See Theorem 8 (g).

<sup>64</sup>  $\rightarrow$  denotes uniform convergence.

- (c). For  $i = 1, \dots, k$ ,  $\lambda_i$  is either in  $\rho(T)$  or else a pole of  $R_\lambda(T)$  of order at most  $\nu_i$ .
- (d).  $R\{F(T)\} \oplus N\{F(T)\} = \mathfrak{X}$ ,  $R\{F(T)\}$  is closed.
- (e).  $R\{(T - \lambda_i I)^{\nu_i+1}\}$  is closed,  $i = 1, \dots, k$ .

Furthermore when the limit  $P$  exists, the set  $\sigma = (\lambda_1, \dots, \lambda_k) \cap \sigma(T)$  is a spectral set of  $T$  and  $P = \frac{1}{2\pi i} \int_C R_\lambda(T) d\lambda$ , where  $C$  is the boundary of some  $D = D(T)$ , such that<sup>65</sup>  $\sigma = D \cap \sigma(T)$ .

Our definition of g.i. is based on the following:

COROLLARY 6. For any bounded symmetric linear operator  $A$  on a Hilbert space  $\mathfrak{H}$  into itself, the sequence

$$(110) \quad e^{-A^{2n}} = \frac{1}{2\pi i} \int_C e^{-\lambda^{2n}} R_\lambda(A) d\lambda, \quad n = 1, 2, \dots,$$

(where  $C$  is the boundary of some  $D = D(A)$  which contains  $\sigma(A)$ ) converges uniformly to a projection on  $N(A)$ .

Proof. In Theorem 21 let

$$F(\lambda) = \lambda, \quad f_n(\lambda) = e^{-\lambda^{2n}}, \quad n = 1, 2, \dots.$$

- (i).  $f_n(\lambda) \in \mathfrak{F}(A)$  and  $f_n(0) = 1, n = 1, 2, \dots$ .
- (ii). For  $C$  as in (110),  $n = 1, 2, \dots$ ,

$$F(A)f_n(A) = \frac{1}{2\pi i} \int_C \lambda e^{-\lambda^{2n}} R_\lambda(A) d\lambda \rightarrow 0,$$

since

$$F(\lambda)f_n(\lambda) = \lambda e^{-\lambda^{2n}} \rightarrow 0$$

uniformly in the whole plane.

(iii). Since  $A$  is a bounded symmetric linear operator,

$$(111) \quad \mathfrak{H} = R(A) \oplus N(A),$$

$R(A)$  is closed.

(iv). In (i), (ii) and (iii), we have established (i), (ii) and (d), respectively, of Theorem 21. Thus, in particular, part (a) of Theorem 21 follows, i.e., the sequence (110)  $\rightarrow P_{N(A)}$ .

As the sequence of scalar functions  $f_n(\lambda)$ , written as

$$(112) \quad f_n(\lambda) = e^{-\lambda^{2n}} = 1 - \int_0^n \lambda^2 e^{-\lambda^{2x}} dx, \quad n = 1, 2, \dots,$$

---

<sup>65</sup> Thus  $P$  is the projection  $P_\sigma$  associated with the spectral set  $\sigma$ , e.g. Riesz and Sz. Nagy [38, p. 419].

was shown to induce a sequence of operators  $f_n(A)$  converging uniformly to  $P_{N(A)}$ , we identify the latter with<sup>66</sup>  $(I - A^+A)$ , and look for a sequence  $g_n(\lambda)$  for which  $g_n(A)$  converges, in some sense, to  $A^+$ .

Such a sequence is

$$(113) \quad g_n(\lambda) = \int_0^n \lambda e^{-\lambda^2 x} dx, \quad n = 1, 2, \dots,$$

rewritten as

$$(114) \quad \begin{cases} g_n(\lambda) = \frac{1}{\lambda} (1 - e^{-\lambda^2 n}), \\ g_n(0) = 0. \end{cases} \quad \text{for } \lambda \neq 0,$$

Clearly  $g_n(\lambda) \in \mathfrak{F}(A)$ ,  $n = 1, 2, \dots$ . Hence the following:

*Definition 10.* For any bounded symmetric linear operator  $A: \mathfrak{K} \rightarrow \mathfrak{K}$ , and  $g_n(\lambda)$  as in (113), consider the sequence

$$(115) \quad A_n^+ = g_n(A) = \frac{1}{2\pi i} \int_C g_n(\lambda) R_\lambda(A) \cdot d\lambda,$$

where  $C$  is as in (113). The *generalized inverse* of  $A$  is defined as<sup>67</sup>

$$(116) \quad A^+ = \lim_{n \rightarrow \infty} A_n^+.$$

That is,  $A^+$  is defined only for the vectors  $x$  which belong to the domains of every  $A_n^+$ , and for which the sequence  $A_n^+x$  converges.

For such  $x$ ,

$$A^+x = \lim_{n \rightarrow \infty} A_n^+x.$$

As the uniqueness of  $A^+$ , so defined, is evident, we will now establish its existence by showing that its domain is nonempty. In fact  $D(A^+)$  is everywhere dense in  $H$ .

In the uninteresting case where  $A$  is nonsingular,  $A^+$  turns to be  $A^{-1}$ . Since  $A$  is nonsingular  $\Leftrightarrow \sigma(A) \cap S_\epsilon = \emptyset$  for some  $\epsilon$ -sphere  $S_\epsilon = \{\lambda: |\lambda| \leq \epsilon\}$ ,  $\epsilon > 0$ , which in turn  $\Rightarrow g_n(\lambda) \rightarrow \frac{1}{\lambda}$  uniformly in some

$D = D(A)$  which contains  $\sigma(A)$ , and  $\frac{1}{\lambda} \in \mathfrak{F}(A)$ , we conclude

$$\lim_{n \rightarrow \infty} A_n^+ = A^{-1}$$

(where the convergence in definition 10 is uniform).

<sup>66</sup> Compare with Theorem 18.

<sup>67</sup> E.g., Riesz and Sz. Nagy [38, p. 299].

We present now some properties of  $A^+$ , which in the general case is an unbounded operator. These will show in particular, the equivalence between our definition 10 and Tseng's (definition 4) for the bounded symmetric case.

**THEOREM 22.** *Let  $A, A_n^+$  and  $A^+$  be as in definition 10. Then*

- (a).  $N(A) = N(A_n^+) = N(A^+)$ , for  $n = 1, 2, \dots$ .
- (b).  $R(A) \subseteq D(A^+)$ .
- (c).  $\overline{AA^+} = P_{R(A)}$ ,  $A^+$  is closed.
- (d).  $\overline{D(A^+)} = \mathfrak{C}$ .
- (e).  $A^+$  is closed.
- (f).  $A^+$  is self adjoint.
- (g).  $A^+A = \overline{P_{R(A^+)}}$ .

*Proof.* (a). (i). The sequence  $h_n(\lambda) = \int_0^n e^{-\lambda^2 x} dx$  belongs to  $\mathfrak{F}(A)$  for all  $n$ . The operators  $A_n^+$  are bounded, thus  $D(A_n^+) = \mathfrak{C}$ .

For every  $x, y \in \mathfrak{C}, n = 1, 2, \dots$ , the following holds<sup>68</sup>:

$$\begin{aligned} (A_n^+ x, y) &= \int_m^M \left\{ \int_0^n \lambda e^{-\lambda^2 x} dx \right\} d(E_\lambda x, y) \\ &= \int_m^M \lambda h_n(\lambda) d(E_\lambda x, y) = \int_m^M h_n(\lambda) d(E_\lambda Ax, y). \end{aligned}$$

Therefore  $N(A) = N(A_n^+), n = 1, 2, \dots$ .

(ii). Since  $A_n^+$  is bounded,  $N(A_n^+)$  is closed for all  $n$ . From definition 10 we conclude  $N(A_n^+) = N(A^+), n = 1, 2, \dots$ .

(b).  $y \in R(A) \Leftrightarrow y = Ax$  for some  $x \in \mathfrak{C}$ .

$$A_n^+ y = A_n^+ Ax = \{I - (I - A_n^+ A)\}x \rightarrow (I - P_{N(A)})x,$$

by Corollary 6. Therefore  $y \in D(A^+)$  (see definition 10).

(c). (i). By (111) and the proof of (b),  $AA^+Ax = AP_{R(A)}x = Ax$ , for all  $x \in \mathfrak{C}$ . Therefore  $P_{R(A)} \subseteq AA^+$ .

(ii). From (a),  $x \in N(A) \Rightarrow AA^+x = 0$ . Therefore  $AA^+ = P_{R(A)}$ .

(d). By definition 10,  $D(A^+) = \{x : x \in \bigcap_{n=1}^\infty D(A_n^+), \text{ there exists } \lim_{n \rightarrow \infty} A_n^+ x\}$ . Suppose there exists  $x$  such that  $(x, y) = 0$  for all  $y \in \overline{D(A^+)}$ , in particular, by (b),  $(x, y) = 0$  for all  $y \in R(A)$ . Since  $A$  is symmetric,  $(Ax, z) = 0$  for all  $z \in \mathfrak{C}$ . Therefore  $x \in N(A)$ . By (a),  $x \in D(A^+)$  therefore  $x = 0$ , which proves  $\overline{D(A^+)} = \mathfrak{C}$ .

<sup>68</sup> Here  $m, M$  are respectively the greatest lower and the least upper bounds of  $A$ , e.g. Riesz and Sz. Nagy [38, p. 262].

(e). For any sequence  $x_n \in D(A^+)$  converging to <sup>69</sup>  $x$ ,

$$\lim_{n \rightarrow \infty} A^+ x_n = \lim_{n \rightarrow \infty} (\lim_{m \rightarrow \infty} A_m^+) x_n = \lim_{m \rightarrow \infty} (\lim_{n \rightarrow \infty} A_m^+ x_n) = \lim_{m \rightarrow \infty} A_m^+ x.$$

Therefore,

$$\lim_{n \rightarrow \infty} A^+ x_n = y \Rightarrow x \in D(A^+) \quad \text{and} \quad A^+ x = y.$$

(f). (i). Because of (d) and (e),  $A^{+*}$  exists and  $\overline{D(A^{+*})} = \mathfrak{C}$ .

(ii). Clearly,  $A_n^+$  is symmetric for all  $n$ . Therefore, for any  $x, y \in D(A^+)$ ,

$$(A^+ x, y) = \lim_{n \rightarrow \infty} (A_n^+ x, y) = \lim_{n \rightarrow \infty} (x, A_n^+ y) = (x, A^+ y).$$

Thus  $A^+$  is symmetric, i.e.,  $A^+ \subseteq A^{+*}$ .

(iii). To show that  $A^+$  is self-adjoint, i.e.  $A^+ = A^{+*}$ , is equivalent to showing that  $\sigma(A^+)$  is confined to the real axis. Since  $A_n^+$  is bounded symmetric, for all  $n$ ,

$$\text{Im} \{ \lambda \} \neq 0 \Rightarrow \lambda \in \rho(A_n^+), \quad n = 1, 2, \dots$$

Let  $x$  be any vector in  $\mathfrak{C}$ ,  $\lambda$  nonreal and

$$y_n = (\lambda I - A_n^+)^{-1} x = R_\lambda(A_n^+) x.$$

Now

$$\| R_\lambda(A_n^+) - R_\lambda(A_m^+) \| = \max \left| \frac{1}{\lambda - g_n(\mu)} - \frac{1}{\lambda - g_m(\mu)} \right|$$

with respect to <sup>70</sup>  $\{E_\mu\}$ . Therefore

$$\| R_\lambda(A_n^+) - R_\lambda(A_m^+) \| = \max_{w \cdot r \cdot t \{E_\mu\}} \left| \frac{g_n(\mu) - g_m(\mu)}{(\lambda - g_n(\mu))(\lambda - g_m(\mu))} \right|.$$

Thus

$$\lim_{m, n \rightarrow \infty} \| R_\lambda(A_n^+) - R_\lambda(A_m^+) \| = 0,$$

and hence the sequence  $y_n$  converges to  $y = R_\lambda(A^+) x$ . Therefore

$$\text{Im} \{ \lambda \} \neq 0 \Rightarrow \lambda \in \rho(A^+)$$

which proves  $A^+ = A^{+*}$ .

(g). Because of (f),  $\overline{R(A^+)} = O(N(A^+))$ . Therefore, using (a) and  $(I - A^+ A) = P_{N(A)}$  we verify that

$$A^+ A = P_{\overline{R(A^+)}}.$$

<sup>69</sup> By Moore's theorem on interchanging limits, e.g., Dunford and Schwartz [15, p. 28, lemma 6].

<sup>70</sup> E.g., Riesz and Sz. Nagy [38, p. 349].

*Remarks.* (i). The constructive definition 10 can be extended to the case of an unbounded self-adjoint operator  $A : \mathcal{H} \rightarrow \mathcal{H}$ , by using a representation theorem for such operators due to Riesz and Lorch<sup>71</sup> [39]. This theorem states the decomposition

$$(117) \quad \mathcal{H} = \sum_{n=1}^{\infty} \oplus \mathcal{H}_n$$

of the space  $\mathcal{H}$  into a vector sum of orthogonal subspaces  $\mathcal{H}_n$ , and the decomposition

$$(118) \quad A = \sum_{n=1}^{\infty} A_{(n)}$$

of the unbounded self-adjoint operator  $A : \mathcal{H} \rightarrow \mathcal{H}$  as the denumerable sum of bounded symmetric operators<sup>72</sup>  $A_{(n)} : \mathcal{H}_n \rightarrow \mathcal{H}_n$ . Here the natural definition of the g.i. is

$$(119) \quad A^+ = \sum_{n=1}^{\infty} A_{(n)}^+$$

where the  $A_{(n)}^+$  are defined as in definition 10 (each  $A_{(n)}$  being bounded and symmetric). Statements analogous to Theorem 22 can be proved, with the exception that the projections  $AA^+$ ,  $A^+A$  need not be closed, and must be extended to give the respective projections  $P_{\overline{R(A)}}$  and  $P_{\overline{R(A^+)}}$ .

(ii). A further extension, to the case of an unbounded closed operator  $A : \mathcal{H} \rightarrow \mathcal{H}$  with  $\overline{D(A)} = \mathcal{H}$ , is possible by using the polar representation theorem of Von Neumann [49, p. 307, Theorem 7]. This theorem represents an unbounded closed  $A : \mathcal{H} \rightarrow \mathcal{H}$  with  $\overline{D(A)} = \mathcal{H}$  as

$$(120) \quad A = BW,$$

where

$$(121) \quad B = \sqrt{AA^*}$$

is self-adjoint and nonnegative definite.  $W$  is a partial isometry, suitably normalized to assure uniqueness.

In this case the g.i. is defined as<sup>73</sup>

$$(122) \quad A^+ = W^*B^+,$$

where  $B^+$  is defined by the preceding remark. Analogously to Theorem 22, we can verify here that: (a)  $N(A^+) = N(A^*)$ , (b)  $R(A) \subseteq D(A^+)$ , (c)  $D(A^+) = \mathcal{H}$ , (d)  $R(A^+) = R(A^*)$  and (e)  $A^+$  is closed. As in the

<sup>71</sup> See also Riesz and Sz. Nagy [38, pp. 313-320].

<sup>72</sup> I.e.,  $P_{\mathcal{H}_n}A_{(n)} = A_{(n)}P_{\mathcal{H}_n}$ .

<sup>73</sup> Compare with (41).

preceding remark here, too,  $AA^+$  and  $A^+A$  are not closed in general and must be extended to give the respective projections  $P_{\overline{R(A)}}$  and  $P_{\overline{R(A^*)}}$ .

## REFERENCES

- [1] M. ALTMAN, *Approximation methods in functional analysis*, Lecture Notes Ma 107C, California Institute of Technology (1959).
- [2] A. BEN-ISRAEL AND A. CHARNES, *Projection properties and Neumann-Euler expansions for the Moore-Penrose inverse of an arbitrary matrix*, ONR Research Memo. No. 40, (Northwestern University, The Technological Institute, Evanston, Illinois) (1961).
- [3] ———, *Generalized inverses and the Bott-Duffin network analysis*, ONR Research Memo. No. 66, (Northwestern University, The Technological Institute, Evanston, Illinois) (1962).
- [4] A. BEN-ISRAEL AND S. J. WERSAN, *A least squares method for computing the generalized inverse of an arbitrary matrix*, ONR Research Memo. No. 61, (Northwestern University, The Technological Institute, Evanston, Illinois) (1962).
- [5] A. BJERHAMMAR, *Application of calculus of matrices to method of least squares; with special reference to geodetic calculations*, Trans. Roy. Inst. Tech. Stockholm, 49 (1951), pp. 1-86.
- [6] ———, *Rectangular reciprocal matrices with special reference to geodetic calculations*, Bull. Geodesique, (1951), pp. 188-220.
- [7] ———, *A Generalized Matrix Algebra*, Kungl. Tekn. Hogsk. Handl., Stockholm, 1958.
- [8] R. BOTT AND R. J. DUFFIN, *On the algebra of networks*, Trans. Amer. Math. Soc., vol. 74, 1 (1953), pp. 99-109.
- [9] A. CHARNES, W. W. COOPER AND G. L. THOMPSON, *Constrained generalized medians and linear programming under uncertainty*, ONR Research Memo. No. 41, (Northwestern University, The Technological Institute, Evanston, Illinois, and Carnegie Institute of Technology, Pittsburgh, Pa.) (1961).
- [10] R. E. CLINE, *On the computation of the generalized inverse  $A^+$ , of an arbitrary matrix  $A$ , and the use of certain associated eigenvectors in solving the allocation problem*, Preliminary Report, Statistical and Computing Laboratory, Purdue University, Lafayette, Indiana (1958).
- [11] G. G. DEN BROEDER, JR. AND A. CHARNES, *Contributions to the theory of a generalized inverse for matrices*, Purdue University, Lafayette, Indiana (1957), Republished as ONR Research Memo. No. 39, (Northwestern University, The Technological Institute, Evanston, Illinois) (1962).
- [12] B. A. DENT AND A. NEWHOUSE, *Polynomials orthogonal over discrete domains*, SIAM Rev., 1 (1959), pp. 55-59.
- [13] M. P. DRAZIN, *Pseudo inverses in associative rings and semigroups*, Amer. Math. Monthly, 65 (1958), pp. 506-514.
- [14] N. DUNFORD, *Spectral theory. I: Convergence to projections*, Trans. Amer. Math. Soc., 54 (1943), pp. 185-217.
- [15] N. DUNFORD AND J. SCHWARTZ, *Linear Operators*, Vol. I, Interscience, New York, 1958.
- [16] V. FADDEEVA, *Computational Methods of Linear Algebra*, Dover, New York, 1959.
- [17] L. FANTAPPiE, *Le calcul des matrices*, C.R. Acad. Sci. Paris, 186 (1928), pp. 619-621.

- [18] J. J. FLORENTIN, *Optimal control of continuous time, Markov, stochastic systems*, J. Electronics Control, 10 (1961), pp. 473-488.
- [19] T. N. E. GREVILLE, *The pseudo inverse of a rectangular or singular matrix and its applications to the solution of systems of linear equations*, SIAM Rev., 1 (1959), pp. 38-43.
- [20] ———, *Some applications of the pseudo inverse of a matrix*, SIAM Rev., 2 (1960), pp. 15-22.
- [21] M. R. HESTENES, *Inversion of matrices by biorthogonalization and related results*, this Journal, 6 (1958), p. 84.
- [22] ———, *Relative hermitian matrices*, Pacific J. Math., 1 (1961), pp. 225-245.
- [23] E. HILLE AND R. S. PHILLIPS, *Functional Analysis and Semi-Groups*, Amer. Math. Soc. Colloq. Publications, Vol. XXXI, Revised edition, New York, 1957.
- [24] R. E. KALMAN, *Contributions to the theory of optimal control*, Bol. Soc. Mat. Mexicana (1960), pp. 102-119.
- [25] ———, *New methods and results in linear prediction and filtering theory*, Technical Report No. 61-1 RIAS, Baltimore, Md. (1961).
- [26] A. T. LONSETH, *Approximate solution of Fredholm type integral equations*, Bull. Amer. Math. Soc., 60 (1954), pp. 415-430.
- [27] C. C. MACDUFFEE, *The Theory of Matrices*, Springer Verlag, Berlin, 1933.
- [28] N. H. MCCOY, *Rings and Ideals*, The Carus Mathematical Monographs, No. 8, Amer. Math. Assoc., Buffalo, 1948.
- [29] E. H. MOORE, ———, Bull. Amer. Math. Soc., 26 (1920), pp. 394-395.
- [30] ———, *General Analysis, Part I*, Mem. Amer. Philos. Soc., I, 1935.
- [31] W. D. MUNN AND R. PENROSE, *A note on inverse semigroups*, Proc. Cambridge Philos. Soc., 51 (1955), pp. 396-399.
- [32] F. J. MURRAY AND J. VON NEUMANN, *On rings of operators-I*, Ann. of Math., 37 (1936), pp. 116-229.
- [33] M. A. NAIMARK, *Normed Rings*, P. Noordhoff N.V., Groningen, The Netherlands, 1959.
- [34] R. PENROSE, *A generalized inverse for matrices*, Proc. Cambridge Philos. Soc., 51 (1955), pp. 406-413.
- [35] ———, *On best approximate solution of linear matrix equations*, Proc. Cambridge Philos. Soc., 52 (1956), pp. 17-19.
- [36] L. D. PYLE, *A gradient projection method for solving programming problems using the generalized inverse  $A^+$  and the eigenvectors  $e(1), \dots, e(n)$  of  $I - A^+A$* , Preliminary Report, Statistical and Computing Laboratory, Purdue University, Lafayette, Indiana (1958).
- [37] R. RADO, *Note on generalized inverses of matrices*, Proc. Cambridge Philos. Soc., 52 (1960), pp. 600-601.
- [38] F. RIESZ AND B. SZ. NAGY, *Functional Analysis*, Ungar, New York, 1955.
- [39] F. RIESZ AND E. LORCH, *The integral representation of unbounded self adjoint transformations in Hilbert space*, Trans. Amer. Math. Soc., 39 (1936), pp. 331-340.
- [40] R. F. RINEHART, *The equivalence of definitions of a matrix function*, Amer. Math. Monthly, 62 (1955), pp. 395-414.
- [41] J. B. ROSEN, *The gradient projection method for nonlinear programming. Part I: Linear constraints*, this Journal, 8 (1960), pp. 181-217.
- [42] ———, *The gradient projection method for nonlinear programming. Part II: Non-linear constraints*, this Journal, 9 (1961), pp. 514-532.

- [43] H. SCHWERTFFEGER, *Les Fonctions de Matrices*, Actualites Sci. Ind., Paris, 1938.
- [44] YU. YA. TSENG, *The characteristic value problem of Hermitian functional operators in a non-Hilbertian space*, Ph.D. thesis, University of Chicago, 1933 (A private edition was published by the University of Chicago Libraries, 1936).
- [45] ———, *Sur les solutions des equations operatrices fonctionnelles entre les espaces unitaires. Solutions extremales. Solutions virtuelles*, C.R. Acad. Sci. Paris, 228 (1949), pp. 640–641.
- [46] ———, *Generalized inverses of unbounded operators between two unitary spaces*, Dokl. Akad. Nauk SSSR (N.S.), 67 (1949), pp. 431–434 (Reviewed in Math. Rev., 11 (1950), p. 115).
- [47] ———, *Properties and classification of generalized inverses of closed operators*, Dokl. Akad. Nauk. SSSR (N.S.), 67 (1949), pp. 607–610 (Reviewed in Math. Rev., 11 (1950), p. 115).
- [48] ———, *Virtual solutions and general inversions*, Uspehi. Mat. Nauk (N.S.), 11 (1956), pp. 213–215 (Reviewed in Math. Rev., 18 (1957), p. 749).
- [49] J. VON NEUMANN, *Über adjungierte Funktionaloperatoren*, Ann. of Math., 33 (1932), pp. 294–310.
- [50] ———, *On regular rings*, Proc. Nat. Acad. Sci. U.S.A., 22 (1936), pp. 707–713.
- [51] ———, *Continuous Geometry*, Princeton University Press, Princeton, New Jersey, 1960.
- [52] ———, *Functional operators. Vol. II: The geometry of orthogonal spaces*, Annals of Mathematics Studies No. 29, Princeton University Press, Princeton, New Jersey, 1950.
- [53] J. H. M. WEDDERBURN, *Lectures on Matrices*, Vol. XVIII, Amer. Math. Soc. Colloq. Publications, New York, 1934.
- [54] J. W. GIBBS, *Collected Works*, Vol. II, part 2, Yale University Press, New Haven, 1931, pp. 61–65.
- [55] F. L. HITCHCOCK, *A solution of the linear matrix equation by double multiplication*, Proc. Nat. Acad. Sci., 8 (1922), pp. 78–83.
- [56] J. H. M. WEDDERBURN, *The absolute value of the product of two matrices*, Bull. Amer. Math. Soc., 31 (1925), pp. 304–308.