Generalized inverses for the members of a class of non-linear operators
by Keith Howard Senechal

A thesis submitted to the Graduate Faculty in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY in Mathematics
Montana State University

Abstract:
Until the present, generalized inverses of only linear operators have appeared in the literature. The primary purpose of this thesis is to define and develop properties of a generalized inverse A‡ for each member A of a class of nonlinear operators. Each A is the restriction of a linear operator a to a translation Zo+D of a linear subset D of the domain X of &alpha:. The operator a maps X into Y where X and Y are linear inner product spaces.

The domain of A † is taken as that subset of Y for which a particular operator system is solvable. The value of A‡, corresponding to each element of its domain, is equal to the unique solution of the operator system. A minimization property of A‡ is also discussed.

GENERALIZED INVERSES FOR THE MEMBERS OF A CLASS OF NON-LINEAR OPERATORS

by

KEITH HOWARD SENECHAL

A thesis submitted to the Graduate Faculty in partial fulfillment of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Mathematics

Approved:

Robert D. Engle
Head, Major Department

Louis C. Barrett
Chairman, Examining Committee

Henry L Parsons
Graduate Dean

MONTANA STATE UNIVERSITY
Bozeman, Montana

August, 1973

ACKNOWLEDGEMENT

The author would like to express his gratitude to Professor Louis C. Barrett for his guidance in researching this thesis topic and his invaluable aid throughout the writing of this thesis.

TABLE OF CONTENTS

ABSTRACT

Until the present, generalized inverses of only linear operators have appeared in the literature. The primary purpose of this thesis is to define and develop properties of a generalized inverse $A^{\ddagger}$ for each member A of a class of nonlinear operators. Each A is the restriction of a linear operator $\mathcal{a}$ to a translation z_0+D of a linear subset D of the domain X of $\mathcal{a}$. The operator $\mathcal{a}$ maps X into Y where X and Y are linear inner product spaces.

The domain of $A^{\ddagger}$ is taken as that subset of Y for which a particular operator system is solvable. The value of $A^{\ddagger}$, corresponding to each element of its domain, is equal to the unique solution of the operator system. A minimization property of $A^{\ddagger}$ is also discussed.

INTRODUCTION

A unique generalized inverse for an arbitrary m by n matrix was first introduced by E. H. Moore in a presentation entitled "On the reciprocal of the general algebraic matrix" at the fourteenth western meeting of the American Mathematical Society in April 1920.

Some thirty odd years later, Penrose [8] defined a generalized inverse for an arbitrary m by n matrix A as the unique solution $X = A^{\dagger}$ of the matrix system:

$$\text{(i)} \qquad AXA = A$$

$$\text{(ii)} \qquad XAX = X$$

$$\text{(iii)} \qquad (AX)^{*} = AX$$

$$\text{(iv)} \qquad (XA)^{*} = XA$$

where * indicates conjugate transpose. He also proved a theorem which states: A necessary and sufficient condition for the matrix equation $AXB = C$ to have a solution is that $AA^{\dagger}CB^{\dagger}B = C$ shall hold. Then, the solutions are $X = A^{\dagger}CB^{\dagger} + (Y - A^{\dagger}AYBB^{\dagger})$ where Y is an arbitrary matrix having the same number of rows and columns as X.

Penrose [9] used his generalized inverse to find least square approximate solutions of inconsistent systems of linear equations. Much of Penrose's work has been carried over to infinite dimensional matrices by Ben-

Israel and Wersan [2].

Tseng [11] extended the concept of a generalized inverse to densely defined linear operators between Hilbert spaces. A paper by Ben-Israel and Charnes [1] gives a good history and bibliography of the subject prior to the year 1963. Subsequent articles by Bradley [3], Loud [6,7], Reid [10], and Krall [4,5] indicate the current state of the theory.

So far, generalized inverses of only linear operators have appeared in the literature. In each instance, the domain of the linear operator has been taken as a dense subset, and the range of the operator as a closed subset, of some complete linear inner product space(s). Some authors have used differential operators defined on such spaces to obtain their results.

The inverse operator considered in this thesis is more general. For neither the operator A, whose inverse is desired, nor its generalized inverse $A^{\ddagger}$, are required to be linear. Instead, the domain of A is allowed to be the translation z_0+D of some linear subset D of a linear inner product space X. Thus, the domain of A need not be linear, nor is it required to be dense in X. On the other hand, we do demand that A be the restriction of a linear operator $\mathcal{A}$ to z_0+D. As for the range of A, we do not insist that it be closed, but merely that it shall be a subset of some linear inner product space which may or may not be complete.

Thus, our generalized inverse is defined for the members of a class of operators whose domains may be translations of a linear subset of a linear inner product space and whose ranges need not be closed. When dealing with an operator whose domain is dense, we may freely use the corresponding adjoint operator just as other writers have done.

The subject matter of this thesis is divided into six chapters and two appendices. Chapter I contains several important theorems concerning properties of sets, and culminates with a discussion of the projection operator.

In Chapter II the adjoint operator is defined and proven to be linear. Our treatment of the projection and adjoint operators is somewhat non-standard. For this reason, we have devoted separate chapters to each of them.

The third chapter deals primarily with solution sets, and conditions for solvability, of a class of linear operator systems. Necessary and sufficient conditions for the existence of solutions are found. A kind of "best" approximate solution for one of these systems is also described and characterized in terms of the solution set of a related system.

Chapter IV has to do with the generalized inverse operator. Once the operator has been defined, we prove that it is usually non-linear. A number of its most important properties are also established.

To illustrate the ideas developed in this thesis, we have added Chapter V. It contains a boundary value problem which serves as an illustrative example. This problem is kept rather simple so as to render the concepts involved quite transparent.

Topics for future research are outlined in Chapter VI. Appendix A includes some standard results pertaining to the behavior of different sets. Finally, Appendix B has been added to render this study more nearly self-contained. The appendices are designed to provide a lucid review of the fundamental notions and structure associated with linear inner product spaces. They also clarify much of the notation employed.

CHAPTER I

THE PROJECTION OPERATOR

The projection operator will be encountered repeatedly in the ensuing work. To facilitate a discussion of such operators, we first state and prove several theorems having to do with a linear subset D of a linear inner product space X. For each element d of D and x of X, the norm $\|x-d\|$ has 0 as a lower bound. In particular, if x is fixed, $\|x-d\|$ has a greatest lower bound

$$b = \operatorname{glb}_{d \in D} \|x-d\| .$$

THEOREM 1.1: Let M be a closed linear subset of a complete linear inner product space X and let x denote any fixed element of X. Set $b = \operatorname{glb}_{d \in D} \|x-m\|$. Then the condition $b = \|x-m\|$ is satisfied by a unique element m_0 of M.

PROOF: In virtue of the definition of b, the statement

$$\forall \text{ pos int } n \; \exists m \in M \; [\, \|x-m\| < b + 1/n \,]$$

must hold. Otherwise, for some n, $b + 1/n$ would be a lower bound of $\|x-m\|$ greater than b. Using the axiom of choice, we form a sequence $\{m_n\}$ in M such that

$\| x-m_n \| < b + 1/n$, and proceed to prove it is a Cauchy sequence. To do this, we must show that

$$\forall \epsilon > 0 \ \exists \text{ pos int } n_0 \ \forall j, k > n_0 [\ \| m_j - m_k \| < \epsilon\ ]$$

is true.

For arbitrary positive integers j and k we have

$$\| m_j - m_k \|^2 = \| (x-m_k) - (x-m_j) \|^2$$
$$= \| x-m_k \|^2 - 2\mathrm{Re}(x-m_k, x-m_j) + \| x-m_j \|^2.$$

Moreover,

$$4 \| x - \tfrac{1}{2}(m_j + m_k) \|^2 = \| (x-m_k) + (x-m_j) \|^2$$
$$= \| x-m_k \|^2 + 2\mathrm{Re}(x-m_k, x-m_j) + \| x-m_j \|^2.$$

Eliminating $2\mathrm{Re}(x-m_k, x-m_j)$ between these two results, we obtain

$$\| m_j - m_k \|^2 = 2 \| x-m_j \|^2 + 2 \| x-m_k \|^2 - 4 \| x - \tfrac{1}{2}(m_j + m_k) \|^2.$$

But, $\frac{1}{2}(m_j + m_k)$ is an element of M because M is linear.

Therefore, $$\| x - \tfrac{1}{2}(m_j + m_k) \|^2 \geqq b^2$$

and $$-4 \| x - \tfrac{1}{2}(m_j + m_k) \|^2 \leqq -4b^2.$$

It follows that

$$\| m_j - m_k \|^2 < 2(b + 1/j)^2 + 2(b + 1/k)^2 - 4b^2$$
$$= 4b(1/j + 1/k) + 2(1/j^2 + 1/k^2)$$
$$\leqq 4b(1/j + 1/k) + 4(1/j + 1/k)$$
$$= 4(b+1)(1/j + 1/k).$$

Now, given an ϵ greater than zero, it is readily verified that

$$4(b+1)(1/j + 1/k) < \epsilon^2$$

for all positive integers j and k greater than n_0, where

$$n_0 \geqq 8(b+1)/\epsilon^2.$$

Thus, the statement

$$\forall \epsilon > 0 \; \exists \text{ pos int } n_0 \; \forall j, k > n_0 [\, \| m_j - m_k \| < \epsilon \,]$$

holds and we conclude that $\{m_n\}$ is a Cauchy sequence. Since X is complete, $\{m_n\}$ must converge to some element m_0 of X. On the other hand, m_0 belongs to M because M is closed. Hence, $\| x - m_0 \| \geqq b$.

We next prove that $\| x - m_0 \| = b$. Let $\epsilon > 0$ be given. Since $\{m_n\}$ converges to m_0, there exists a positive integer n_1 such that $\| m_n - m_0 \| < \epsilon/2$ for all n greater than n_1. Choose an integer n_2 which is greater than $\max(2/\epsilon, n_1)$. Then, for every $n \geqq n_2$,

$$\begin{aligned} \| x - m_0 \| &\leqq \| x - m_n \| + \| m_n - m_0 \| \\ &< b + 1/n + \epsilon/2 \\ &< b + \epsilon/2 + \epsilon/2 = b + \epsilon. \end{aligned}$$

Since $\| x - m_0 \|$ is independent of ϵ, we conclude that $\| x - m_0 \| \leqq b$. But, $\| x - m_0 \| \geqq b$ and therefore, $\| x - m_0 \| = b$. Thus, there is at least one element of M, namely m_0, which satisfies the condition $\| x - m \| = b$.

To complete the proof of the theorem, we shall show that m_0 is the only member of M that fulfills this condition. Suppose r belongs to M and $\|x-r\| = b$. By steps that are already familiar, we find

$$\|m_o-r\|^2 = \|(x-r) - (x-m_o)\|^2$$
$$= 2\|x-r\|^2 + 2\|x-m_o\|^2 - 4\|x-\tfrac{1}{2}(r+m_o)\|^2$$
$$= 2b^2 + 2b^2 - 4\|x-\tfrac{1}{2}(r+m_o)\|^2.$$

Now, $\frac{1}{2}(r+m_o)$ is an element of M, so $\|x-\frac{1}{2}(r+m_o)\|^2 \geqq b^2$. It follows that $\|m_o-r\| = 0$. Denoting the zero element of X by Φ, we conclude that $m_o-r = \Phi$, and $m_o = r$.∎

<u>THEOREM 1.2</u>: Let $M^\perp$ be the orthogonal complement of a closed linear subset M of a complete linear inner product space X. Then

$$M \oplus M^\perp = X.$$

<u>PROOF</u>: Since M and $M^\perp$ are mutually orthogonal linear subsets of X, and X is linear, $M \oplus M^\perp$ is a subset of X. Thus, $M \oplus M^\perp = X$ will hold if $X \subset M \oplus M^\perp$ is true. To establish the latter relation, consider an arbitrary element x of X. With x fixed, there is a single element m_o of M which minimizes $\|x-m\|$ over M. This follows from the fact that M is closed, and X is complete, so that Theorem 1.1 applies. Express x as the sum $m_o + (x-m_o)$ and set $b = \|x-m_o\|$. Since m_o is in M, our proof will be

complete if we can show that $x-m_0$ belongs to $M^{\perp}$; for every element of X will then be a sum $m+p$ of an element m in M and an element p in $M^{\perp}$. We proceed to prove the proposition

$$\forall m \in M[(x-m_0, m) = 0].$$

Let m be an arbitrary element of M. If $m = \Phi$, then $(x-m_0, m) = 0$. Assume m does not equal Φ and let α be an arbitrary element of the field. Then,

$$b^2 \leqq \| x-(m_0+\alpha m) \|^2.$$

Using the defining and antilinearity properties of the inner product, we obtain

$$\begin{aligned} b^2 &\leqq \| (x-m_0)-\alpha m \|^2 \\ &= \| x-m_0 \|^2 - 2\mathrm{Re}(x-m_0, \alpha m) + \| \alpha m \|^2 \\ &= b^2 - 2\mathrm{Re}[\overline{\alpha}(x-m_0, m)] + |\alpha|^2 \| m \|^2. \end{aligned}$$

Setting $\alpha = (x-m_0, m)/\| m \|^2$, it follows that

$$|(x-m_0, m)|^2 \leqq 0;$$

consequently, $(x-m_0, m) = 0$. Therefore, $x-m_0$ is an element of $M^{\perp}$ so that X is a subset of $M \oplus M^{\perp}$.‖

<u>THOEREM</u> <u>1.3</u>: Let D be a linear subset of a complete linear inner product space X. Then D is dense in X if and only if Φ is the only element of X orthogonal to D.

PROOF: Let D be dense in X and let x_0 be an element of X which is orthogonal to D. We first show that x_0 equals Φ. To do this, we choose an arbitrary ϵ greater than zero. Due to the denseness of D in X, there is a d in D such that $\| x_0 - d \| < \epsilon$. In virtue of the linearity of the inner product, and the orthogonality property $(x_0, d) = 0$, the following relations are true.

$$\begin{aligned} \| x_0 - d \|^2 &= \| x_0 \|^2 - 2\mathrm{Re}(x_0, d) + \| d \|^2 \\ &= \| x_0 \|^2 + \| d \|^2 \\ &\geqq \| x_0 \|^2 . \end{aligned}$$

Since ϵ may be taken arbitrarily small, we conclude that $\| x_0 \|^2 = 0$, because x_0 is independent of ϵ. Therefore, $x_0 = \Phi$. (Notice that this part of the proof does not require D to be linear nor that X be complete.)

Now, assume that Φ is the only element of X that is orthogonal to D. To establish that D is dense in X, it will suffice to prove that $\overline{D} \oplus \{\Phi\} = X$; for then, $\overline{D} = X$, where $\overline{D}$ denotes the closure of D. As an initial step note that $\overline{D}$ is linear since the closure of a linear subset is itself a linear subset. (Theorem A.1)

Since $\overline{D}$ is closed and linear, we know from Theorem 1.2 that $X = \overline{D} \oplus \overline{D}^{\perp}$. The set $\overline{D}^{\perp}$ contains Φ. To more closely examine this set, suppose it contains p. Then p is orthogonal to D, because D is a subset of $\overline{D}$. But, by our original assumption, Φ is the only element of X orthogonal to D. Therefore, $p = \Phi$, $\overline{D}^{\perp} = \{\Phi\}$, and $X = \overline{D} \oplus \{\Phi\}$. However, $\overline{D} \oplus \{\Phi\} = \overline{D}$ so that $X = \overline{D}$. We conclude that D is dense in X.∎

<u>COROLLARY 1.3</u>: If D is dense in a linear inner product space X, the only element of X orthogonal to D is Φ.

<u>PROOF</u>: See the parenthetical remark to the first part of the preceding proof.∎

<u>THEOREM 1.4</u>: Let D be a linear subset of a complete linear inner product space X. Then $D \oplus D^{\perp}$ is dense in X.

<u>PROOF</u>: Since D is linear, so is $D \oplus D^{\perp}$. Were D not linear, $D \oplus D^{\perp}$ would not be defined (Appendix B). Assume some element x of X is orthogonal to $D \oplus D^{\perp}$. Then x is orthogonal to D, because D is a subset of $D \oplus D^{\perp}$. Consequently, x is an element of $D^{\perp}$. On the other hand, x is orthogonal to $D^{\perp}$, because $D^{\perp}$ is also a subset of $D \oplus D^{\perp}$. Since x is both in and orthogonal to $D^{\perp}$, this means that x must be orthogonal to itself. But, $(x,x) = 0$ if and only if $x = \Phi$.

We have shown that the only element of X that is orthogonal to $D \oplus D^{\perp}$ is Φ. From Theorem 1.3, we conclude that $D \oplus D^{\perp}$ is dense in X.∎

We note in passing, that an element can be orthogonal to $D^{\perp}$ without necessarily being in D. However, if D is closed and linear, then $(D^{\perp})^{\perp} = D$ (Theorem A.2).

<u>THEOREM</u> <u>1.5</u>: Let D be a linear subset of a linear inner product space X. Suppose x is any element of $D \oplus D^{\perp}$. Then there is a unique element d_o of D such that $x-d_o$ is an element of $D^{\perp}$.

An equivalent way of stating the conclusion of this theorem is that any given element x of $D \oplus D^{\perp}$ can be written as a sum $x = d_o + (x-d_o)$ of elements d_o in D and $x-d_o$ in $D^{\perp}$, in just one way. Thus, $x-d_o$, as well as d_o, is unique.

<u>PROOF</u>: Since x is an element of $D \oplus D^{\perp}$, it can be written as the sum of an element d_o of D, and an element of $D^{\perp}$ which must be $x-d_o$. Assume x can also be expressed as the sum $x = d_1 + (x-d_1)$ of elements d_1 in D and $x-d_1$ in

$D^{\perp}$. Since $D^{\perp}$ is linear, $(x-d_0) - (x-d_1) = d_1 - d_0$ is in $D^{\perp}$. On the other hand, $d_1 - d_0$ is an element of D. This implies that $d_1 - d_0 = \Phi$ and finally that $d_1 = d_0$.∎

With this theorem established, we are now prepared to introduce what will be called the projection operator $P_D: D \oplus D^{\perp} \rightarrow D$, having domain $D \oplus D^{\perp}$ and range D.

DEFINITION: Let X be a linear inner product space and let D be a linear subset of X. For each x in the set $D \oplus D^{\perp}$, define $P_D x$ to be that unique element d_0 in D for which $x - d_0$ is an element of $D^{\perp}$. The operator P_D, thus defined, will be called the projection operator from $D \oplus D^{\perp}$ onto D. The element d_0 in D, assigned to a given x in $D \oplus D^{\perp}$ by P_D, is said to be the projection of x onto D.

Note that Theorem 1.4 ensures that the operator P_D is densely defined whenever X is a complete space; that is to say, the domain $D \oplus D^{\perp}$ of P_D is dense in X. Theorem 1.2 establishes that the operator is defined on all of X if, in addition to X being complete, D is closed.

THEOREM 1.6: The projection operator P_D, has the following properties:

(a) $\forall x \in D \oplus D^{\perp} [x \in D \text{ iff } P_D x = x]$

(b) $\forall x \in D \oplus D^{\perp} [x \in D^{\perp} \text{ iff } P_D x = \Phi]$

(c) $\forall x \in D \oplus D^{\perp} [x - P_D x \in D^{\perp}]$

(d) $\forall x \in D \oplus D^{\perp} [P_D(P_D x) = P_D x]$

(e) $\forall x \in D \oplus D^{\perp} [P_{\overline{D}} x = P_D x]$

(f) P_D is a linear operator

Statements (a), (b), and (c) follow immediately from the definition of P_D. Since $P_D x$ is in D, statement (d) is a direct consequence of (a). As for (e) and (f), they require explicit verification.

PROOF(e): To establish property (e), we shall prove that P_D is the restriction of $P_{\overline{D}}$ to $D \oplus D^{\perp}$. Since $P_{\overline{D}}$ has domain $\overline{D} \oplus \overline{D}^{\perp}$, this means both

$$D \oplus D^{\perp} \subset \overline{D} \oplus \overline{D}^{\perp} \quad \text{and} \quad \forall x \in D \oplus D^{\perp} [P_{\overline{D}} x = P_D x]$$

must be confirmed.

Let x denote an arbitrary element of $D \oplus D^{\perp}$. Express x as the unique sum $x = P_D x + (x - P_D x)$. Since D is a subset of $\overline{D}$, $P_D x$ is in $\overline{D}$. Given any element c_0 of $\overline{D}$, there exists a sequence $\{d_n\}$ of points in D which converges to

c_0. Now, $x-P_Dx$ is in $D^\perp$; hence, for every positive integer n, $(x-P_Dx, d_n) = 0$. But,

$$\begin{aligned}(x-P_Dx, c_0) &= (x-P_Dx, \lim_{n\to\infty} d_n)\\ &= \lim_{n\to\infty} (x-P_Dx, d_n)\\ &= 0\end{aligned}$$

due to the continuity of the inner product. Therefore $x-P_Dx$, being orthogonal to $\overline{D}$, is contained in $\overline{D}^\perp$. We conclude that every element x of $D\oplus D^\perp$ belongs to $\overline{D}\oplus\overline{D}^\perp$, that is to say, $D\oplus D^\perp$ is a subset of $\overline{D}\oplus\overline{D}^\perp$.

In view of this conclusion, any element x of $D\oplus D^\perp$ is given by $P_{\overline{D}}x + (x-P_{\overline{D}}x)$, as well as by $P_Dx + (x-P_Dx)$. Thus, the element

$$(x-P_{\overline{D}}x) - (x-P_Dx) = P_Dx - P_{\overline{D}}x$$

is in the intersection of $\overline{D}^\perp$ and $\overline{D}$. From this, the truth of

$$\forall x \in D\oplus D^\perp [P_{\overline{D}}x = P_Dx]$$

is apparent.∎

<u>PROOF</u>(f): Let x_1 and x_2 be elements of $D\oplus D^\perp$, and let α and β be arbitrary elements of the field. Since $D\oplus D^\perp$ is already known to be linear, in order to prove that P_D is linear, we only need to show that

$$P_D(\alpha x_1 + \beta x_2) = \alpha P_Dx_1 + \beta P_Dx_2 .$$

To this end, write the element $\alpha x_1 + \beta x_2$ as the sum

$$[\alpha P_D x_1 + \beta P_D x_2] + [\alpha x_1 + \beta x_2 - \alpha P_D x_1 - \beta P_D x_2] =$$
$$[\alpha P_D x_1 + \beta P_D x_2] + [\alpha(x_1 - P_D x_1) + \beta(x_2 - P_D x_2)].$$

Since $P_D x_1$ and $P_D x_2$ are elements of D, and D is linear, $\alpha P_D x_1 + \beta P_D x_2$ is in D. According to property (c), $x_1 - P_D x_1$ and $x_2 - P_D x_2$ belong to $D^\perp$. But, $D^\perp$ is also linear. Hence, $\alpha[x_1 - P_D x_1] + \beta[x_2 - P_D x_2]$ is in $D^\perp$. Having succeeded in writing $\alpha x_1 + \beta x_2$ as the unique kind of sum required by the definition of the projection operator, we conclude that

$$P_D(\alpha x_1 + \beta x_2) = \alpha P_D x_1 + \beta P_D x_2 . \blacksquare$$

<u>THEOREM</u> <u>1.7</u>: Let X be a linear inner product space including D as a linear subset. Further, assume x is some fixed element of $D \oplus D^\perp$. Then $d = P_D x$ is the unique element of D which minimizes $\| x-d \|$ over D.

<u>PROOF</u>: With x fixed in $D \oplus D^\perp$, let d denote any element of D. Then, the following equalities hold.

$$\| x-d \|^2 = \| (x-P_D x) + (P_D x-d) \|^2$$
$$= \| x-P_D x \|^2 + 2\mathrm{Re}(x-P_D x, P_D x-d) + \| P_D x-d \|^2 .$$

But, $x-P_D x$ is an element of $D^\perp$ and $P_D x-d$ is an element of D. Therefore, their inner product is zero and

$$\| x-d \|^2 = \| x-P_D x \|^2 + \| P_D x-d \|^2 .$$

Thus, $\| x-d \|$ is a minimum over D if and only if d is the projection of x onto D.∎

Although Theorem 1.7 appears to resemble Theorem 1.1, in that both theorems are concerned with minimizing a norm over a specified subset, it is important to recognize that their hypotheses are quite different. Theorem 1.1 requires M to be a closed linear subset of a complete linear inner product space X. There premises ensure that, for fixed x in X, $\| x-m \|$ is minimized by a unique element m_0 of M. In Theorem 1.7, D is not necessarily closed, nor is X necessarily complete. Yet, for fixed x in $D \oplus D^{\perp}$, $\| x-d \|$ is minimized by a unique element of D, namely the element $P_D x$.

CHAPTER II

THE ADJOINT OPERATOR

Having now gained some familiarity with the projection operator, we next introduce the concept of an adjoint operator.

Let X and Y be linear inner product spaces and denote the inner product of typical elements x_1, x_2 of X and y_1, y_2 of Y by $(x_1, x_2)_X$ and $(y_1, y_2)_Y$, respectively. Let D_A be dense in X and suppose that A is some operator defined on D_A into Y. With each such operator, we associate another operator A*, called the adjoint of A.

To define the adjoint of an operator A, of the type just described, we first specify the domain and then the rule of operation for A*. This operator is to have its domain in Y and its range in X.

Specifically, we take the domain of A* to be the following subset of Y:

$$D_{A*} = \{y \in Y \mid \exists w \in X \; \forall x \in D_A [(Ax, y)_Y = (x, w)_X]\}.$$

This set contains the zero element θ of Y, because the statement

$$\forall x \in D_A [(Ax, \theta)_Y = (x, \Phi)_X]$$

is true. As before, Φ denotes the zero element of X. Hence, D_{A*} is not empty.

We next prove that to each y_o in D_{A*}, there corresponds a unique element w of X which satisfies the condition

$$\forall x \in D_A [(Ax, y_o)_Y = (x, w)_X].$$

To do so, we assume this condition is fulfilled by elements w_1 and w_2 both in X. It follows that

$$\forall x \in D_A [(x, w_1 - w_2)_X = 0]$$

holds due to the antilinearity of the inner product. Therefore, $w_1 - w_2$ is orthogonal to D_A. Recalling that D_A is dense in X, we conclude that $w_1 - w_2 = \Phi$ or that $w_1 = w_2$ from Corollary 1.3. Thus, for every y in D_{A*}, the corresponding w in X is unique.

We now complete the definition of the adjoint operator A* of A by prescribing its rule of operation. Given any y in D_{A*}, we require that A*y shall be the unique element w of X for which the statement

$$\forall x \in D_A [(Ax, y)_Y = (x, w)_X]$$

is true. The useful result,

$$\forall x \in D_A \ \forall y \in D_{A*} [(Ax, y)_Y = (x, A^*y)_X]$$

concerning A and A* then holds.

Although neither the rule nor domain of A is required to be linear, the operator A* is linear as the next theorem demonstrates.

THEOREM 2.1: A^* is a linear operator.

PROOF: First we shall verify that D_{A^*} is a linear subset of Y. To do this, let us suppose y_1 and y_2 are arbitrary elements of D_{A^*}, and that w_1 and w_2 are the corresponding elements of X for which both

$$\forall x \in D_A [(Ax, y_1)_Y = (x, w_1)_X]$$

and

$$\forall x \in D_A [(Ax, y_2)_Y = (x, w_2)_X]$$

hold. Then $A^*y_1 = w_1$ and $A^*y_2 = w_2$. For arbitrary elements α and β of the field, the following equations are satisfied by all x in D_A.

$$\begin{aligned}(Ax, \alpha y_1 + \beta y_2)_Y &= \overline{\alpha}(Ax, y_1)_Y + \overline{\beta}(Ax, y_2)_Y \\ &= \overline{\alpha}(x, w_1)_X + \overline{\beta}(x, w_2)_X \\ &= (x, \alpha w_1 + \beta w_2)_X\end{aligned}$$

Thus, the linear combination $\alpha y_1 + \beta y_2$ is in D_{A^*} and we have shown that D_{A^*} is a linear subset of Y.

In addition, we have shown that

$$\begin{aligned}A^*(\alpha y_1 + \beta y_2) &= \alpha w_1 + \beta w_2 \\ &= \alpha A^* y_1 + \beta A^* y_2 .\end{aligned}$$

Therefore, the rule of operation, as well as the domain, of A^* is linear; thus, A^* is a linear operator.∎

<u>THEOREM 2.2</u>: If X is a complete linear inner product space, then the adjoint of the projection operator onto D is equal to the projection operator onto $\overline{D}$. That is,

$$P_D^* = P_{\overline{D}}.$$

<u>PROOF</u>: Since X is a complete space, Theorem 1.4 ensures that the domain $D \oplus D^{\perp}$ of P_D is dense in X. Thus, it makes sense to talk about the adjoint operator. Before examining the domain ot P_D^*, let us first verify the statement

$$\forall x,y \in X [(P_{\overline{D}}x,y) = (x,P_{\overline{D}}y)].$$

Consider arbitrary elements x and y of X. By Theorem 1.2, $\overline{D} \oplus \overline{D}^{\perp} = X$, so x and y are in the domain of $P_{\overline{D}}$. From Property (c) of Theorem 1.6, with $\overline{D}$ in place of D, $x-P_{\overline{D}}x$ and $y-P_{\overline{D}}y$ are orthogonal to $\overline{D}$, whereas $P_{\overline{D}}x$ and $P_{\overline{D}}y$ are elements of $\overline{D}$. Consequently, it is true that

$$0 = (x-P_{\overline{D}}x,P_{\overline{D}}y) = (x,P_{\overline{D}}y) - (P_{\overline{D}}x,P_{\overline{D}}y)$$

and $$0 = (P_{\overline{D}}x,y-P_{\overline{D}}y) = (P_{\overline{D}}x,y) - (P_{\overline{D}}x,P_{\overline{D}}y).$$

Therefore, $(P_{\overline{D}}x,y) = (x,P_{\overline{D}}y)$ for arbitrary x and y in X.

In view of Property (e) of Theorem 1.6,

$$\forall x \in D \oplus D^{\perp} [P_{\overline{D}}x = P_D x]$$

holds. Hence, restricting x to $D \oplus D^{\perp}$, we have as a special case of our initial symbolic statement

$$\forall y \in X \ \forall x \in D \oplus D^{\perp} [(P_D x,y) = (x,P_{\overline{D}}y)].$$

To finish proving the theorem, we observe that the domain of P_D^* is the set

$$\{y \in X \mid \exists w \in X \ \forall x \in D \oplus D^{\perp} [(P_D x, y) = (x, w)]\}.$$

But, the condition

$$\exists w \in X \ \forall x \in D \oplus D^{\perp} [(P_D x, y) = (x, w)]$$

is satisfied by every y in X with $w = P_{\overline{D}} y$. Thus, the domain of P_D^* is X, which is also the domain of $P_{\overline{D}}$. Furthermore, $P_D^* y = w = P_{\overline{D}} y$ for every y in X. We conclude that $P_D^* = P_{\overline{D}}$ since they have identical domains and identical rules of operation.∎

<u>COROLLARY 2.2</u>: If D is a closed linear subset of a complete linear inner product space X, then $P_D^* = P_D$.

<u>PROOF</u>: In this case $D = \overline{D}$; thus, $D \oplus D^{\perp} = \overline{D} \oplus \overline{D}^{\perp} = X$ and $P_D^* = P_{\overline{D}} = P_D$.∎

When an operator A is the same as its adjoint A*, A is said to be self-adjoint.

CHAPTER III

SOLUTION SETS AND BEST APPROXIMATE SOLUTIONS OF LINEAR OPERATOR SYSTEMS

In this chapter we examine solution sets of linear operator systems. Each system is comprised of a condition, involving a linear operator, and another condition restricting the operator domain to the translation of a linear space. A kind of "best" approximate solution of such a system is also discussed.

By way of notation, let X and Y be linear inner product spaces having Φ and θ as their respective zero elements. In addition, require D to be a linear subset of X. Next, consider a linear operator $\mathcal{Q}:X\rightarrow Y$ and designate the restriction of $\mathcal{Q}$ to D by B, so that $B:D\rightarrow Y$. Since D is linear, and

$$\forall d\in D[\mathcal{Q}d = Bd]$$

is true, B is a linear operator. Hence the range R_B of B, and $R_B \oplus R_B^{\perp}$, are linear. Of course, $R_B^{\perp}$ is also linear and the intersection of R_B and $R_B^{\perp}$ contains only θ. All those elements which B maps onto θ make up a linear subset N_B of D, called the null space of B.

With D prescribed, and for fixed z_o in X, the translation z_o+D of D is a subset of the domain X of $\mathcal{Q}$. Let A designate the restriction of $\mathcal{a}$ to z_o+D. We shall have occasion to use the fact that

$$\forall x \in z_o + D[\mathcal{a}x = Ax].$$

It is to be noticed that A is a linear operator if and only if z_o belongs to D, in which case A is the same as B.

To every condition c(x) having z_o+D as domain, there corresponds a unique solution set comprised of just those members of z_o+D for which c(x) is true. If the solution set is empty, no element of z_o+D satisfies c(x); that is, c(x) has no solution.

Now, let y be any given element of Y. With y and $R_B^{\perp}$ fixed, form the linear operator condition

$$\mathcal{a}x - y \in R_B^{\perp}$$

in the variable x. This condition is defined on X. Upon restricting the domain of this condition to z_o+D, we obtain the non-homogeneous operator system

$$\text{I} \qquad \left[\begin{array}{r} Ax - y \in R_B^{\perp} \\ x \in z_o + D \end{array}\right.$$

This system is the prototype of all the linear operator systems that appear in this chapter. It has a solution if and only if

$$\exists\, x \in z_o + D[Ax - y \in R_B^{\perp}]$$

holds.

For convenience of reference, we shall call $Ax - y \in R_B^\perp$ the operator condition, and $x \in z_0 + D$ the auxiliary condition of System I. If $y \in R_B^\perp$, the operator condition is homogeneous; if $z_0 \in D$, the auxiliary condition is homogeneous. When both $y \in R_B^\perp$ and $z_0 \in D$ hold, as they certainly do when $y = \theta$ and $z_0 = \Phi$, System I reduces to the completely homogeneous system

$$\text{II} \qquad \begin{cases} Bx \in R_B^\perp \\ x \in D \end{cases}$$

With regard to this system, the null space N_B of B is especially important because:

<u>THEOREM</u> <u>3.1</u>: The solution set of System II is N_B.

<u>PROOF</u>: Any solution x_0 of II must belong to D. Hence, Bx_0 is contained in $R_B \cap R_B^\perp = \{\theta\}$. Since B maps x_0 onto θ, every solution of II belongs to N_B.

Conversely, any element x_1 of N_B is in D, whereas $Bx_1 = \theta$ is in $R_B^\perp$. Thus, every element of N_B is a solution of II. We conclude that the solution set of II is N_B. ∎

<u>THEOREM</u> <u>3.2</u>: If System I has a solution x_1, its solution set is $x_1 + N_B$.

PROOF: Let x stand for an arbitrary solution of System I. Inasmuch as x and x_1 are elements of z_0+D, $x-x_1$ is an element of D. Both $Ax - y$ and $Ax_1 - y$ are elements of $R_B^\perp$, and $R_B^\perp$ is linear. Consequently,

$$\begin{aligned}(Ax - y) - (Ax_1 - y) &= \alpha(x - x_1)\\ &= B(x - x_1)\end{aligned}$$

is also an element of $R_B^\perp$. The difference $x-x_1$ is therefore a solution of System II. But, according to Theorem 3.1, $x-x_1$ is a solution of II if and only if $x-x_1 \in N_B$ holds. It follows that $x_1 + N_B$ is the solution set of System I.∎

In case $R_B^\perp$ is replaced by $\{\theta\}$, System I becomes

$$\text{III} \qquad \begin{cases} Ax = y \\ x \in z_0+D \end{cases}$$

THEOREM 3.3: If System III has a solution x_2, its solution set is $x_2 + N_B$.

PROOF: Let x stand for an arbitrary solution of System III. Since x and x_2 are elements of z_0+D, $x-x_2$ is an element of D. Both Ax and Ax_2 are equal to y and α is linear. Consequently,

$$Ax - Ax_2 = \alpha(x - x_2)$$
$$= B(x - x_2)$$
$$= \theta.$$

The difference $x-x_2$ is therefore an element of N_B and x is in the set $x_2 + N_B$.

To complete the proof, suppose that x_2+u is any element of the set $x_2 + N_B$ where u belongs to N_B. Then, x_2+u is an element of $z_o + D$. Furthermore,

$$A(x_2 + u) = \alpha x_2 + \alpha u$$
$$= Ax_2 + Bu$$
$$= y + \theta$$
$$= y.$$

Thus, x_2+u is a solution of System III and $x_2 + N_B$ is the corresponding solution set.∎

<u>THEOREM</u> <u>3.4</u>: If System III has a solution, the solution sets of I and III are identical.

<u>PROOF</u>: Assume x_o is a solution of System III. Then x_o is an element of z_o+D and $Ax_o - y = \theta$ is an element of $R_B^{\perp}$. Therefore, x_o is also a solution of System I. Theorems 3.2 and 3.3 guarantee that $x_o + N_B$ is the solution set of both of the systems I and III.∎

If System I has a solution, we shall usually denote its solution set by $x_o + N_B$.

<u>THEOREM</u> <u>3.5</u>: System III has a solution if and only if $Az_o - y$ is an element of R_B.

<u>PROOF</u>: Suppose System III has the solution x_o. Then x_o is an element of z_o+D and can be written as the sum z_o+d where d is some element of D. We also know that $Ax - y$ equals θ. Thus,

$$\begin{aligned} A(z_o+d) - y &= \alpha(z_o+d) - y \\ &= Az_o - y + Bd \\ &= \theta . \end{aligned}$$

We conclude that

$$\begin{aligned} Az_o - y &= -Bd \\ &= B(-d) \end{aligned}$$

which is an element of R_B.

On the other hand, suppose $Az_o - y$ is an element of R_B. Then there exists an element d of D such that $Az_o - y = Bd$. Thus, $x = z_o - d$ satisfies the equation $Ax = y$ and, furthermore, $x = z_o - d$ is an element of z_o+D. Therefore, System III has a solution, namely, $x = z_o - d$.▮

THEOREM 3.6: System I has a solution if and only if $Az_o - y$ is an element of the set $R_B \oplus R_B^{\perp}$.

PROOF: Assume x_1 is a solution of System I. Then $x_1 = z_o + d$, where d is some element of D. Moreover, the following equalities hold.

$$\begin{aligned} Az_o - y &= A(x_1 - d) - y \\ &= \alpha(x_1 - d) - y \\ &= Ax_1 - y - Bd \\ &= (Ax_1 - y) + B(-d) \end{aligned}$$

But, $Ax_1 - y$ is an element of $R_B^{\perp}$ since x_1 is a solution of System I. Also, $B(-d)$ is an element of R_B. Therefore, $Az_o - y$ is an element of $R_B \oplus R_B^{\perp}$.

To prove the implication in the other direction, assume $Az_o - y$ is an element of $R_B \oplus R_B^{\perp}$. In particular, let $Az_o - y = Bd + v$ where d is an element of D and v is an element of $R_B^{\perp}$. We proceed to show that System I has the solution $z_o - d$.

Owing to the linearity of D, $z_o - d$ is an element of $z_o + D$. What is more,

$$\begin{aligned} A(z_o - d) - y &= \alpha(z_o - d) - y \\ &= (Az_o - y) - Bd \\ &= (Bd + v) - Bd = v \end{aligned}$$

which is an element of $R_B^\perp$. Thus, $z_0 - d$ is a solution of System I.∎

Now, consider the system

$$\text{IV} \qquad \begin{cases} Bx - y \in R_B^\perp \\ x \in D \end{cases}$$

whose auxiliary condition is homogeneous, and the system

$$\text{V} \qquad \begin{cases} Ax \in R_B^\perp \\ x \in z_0 + D \end{cases}$$

having a homogeneous operator condition. Our next theorem reveals that System I is solvable whenever both IV and V are.

<u>THEOREM 3.7</u>: Let Az_0 and y both be elements of $R_B \oplus R_B^\perp$. Then each of Systems IV and V is solvable. Furthermore, if x_1 is any solution of IV and x_2 any solution of V, the solution set of System I is $(x_1 + x_2) + N_B$.

<u>PROOF</u>: Whenever Az_0 and y are both in $R_B \oplus R_B^\perp$, the difference $Az_0 - y$ belongs to $R_B \oplus R_B^\perp$. Then, IV and V have solutions according to Theorem 3.6. Let x_1 be a solution of IV, and x_2 a solution of V. Then

$$\alpha(x_1 + x_2) - y = (Bx_1 - y) + Ax_2$$

is an element of $R_B^{\perp}$. In addition, $x_1 + x_2$ is an element of z_0+D. Thus, $x_1 + x_2$ is a solution of System I. From Theorem 3.2, the solution set of System I is $(x_1 + x_2) + N_B$. ▮

We observe in passing that System I might be solvable even though IV and V are not. For $Az_0 - y$ may be an element of $R_B \oplus R_B^{\perp}$ although Az_0 and y fail to be. If any two of the three systems I, IV, or V are solvable then the third system is solvable also.

In case System III has a solution, its solution set can be expressed in terms of solutions x_1 and x_2 of IV and V, respectively, provided both of these systems are solvable. For, Theorems 3.4 and 3.7 reveal that the solution set of System III is the same as that of System I, namely $(x_1 + x_2) + N_B$. When System III has no solution, but System I does, a "best" approximate solution of System III, in the sense of the following theorem, may be found.

<u>THEOREM</u> <u>3.8</u>: Given an element y in $Az_0 + R_B \oplus R_B^{\perp}$, the condition

$$\forall x \in z_0+D [\ \| Aw - y \| \leq \| Ax - y \|\ ]$$

is satisfied by an element w in z_0+D if and only if w belongs to the solution set $x_0 + N_B$ of System I.

On the other hand,

$$\forall w \in x_0 + N_B [\ \| Aw - y \| = \| Ax - y \|\ ]$$

is satisfied by an element x of z_0+D if and only if x is in $x_0 + N_B$.

PROOF: For each element y of $Az_0 + R_B \oplus R_B^{\perp}$, System I has a solution by Theorem 3.6. Any element w of the solution set $x_0 + N_B$ is in z_0+D. Let x be contained in z_0+D. Then, the following relations hold.

$$\| Ax - y \|^2 = \| A(x-w) + (Aw-y) \|^2$$

$$= \| B(x-w) \|^2 + 2\mathrm{Re}(B(x-w), Aw-y)_Y + \| Aw-y \|^2$$

Since B(x-w) is an element of R_B, while Aw-y is in $R_B^{\perp}$, the inner product $(B(x-w), Aw-y)_Y$ vanishes. Therefore, the statement

$$\forall w \in x_0 + N_B\ \forall x \in z_0 + D [\ \| Ax-y \|^2 = \| B(x-w) \|^2 + \| Aw-y \|^2\ ]$$

is true. From this, we conclude that the condition

$$\forall x \in z_0 + D [\ \| Aw - y \| \leqq \| Ax - y \|\ ]$$

is fulfilled by an element w in z_0+D provided w belongs to the solution set $x_0 + N_B$ of System I.

To prove the converse of the conclusion just stated, suppose w is an element of z_0+D which satisfies the condition $\forall x \in z_0 + D [\ \| Aw - y \| \leqq \| Ax - y \|\ ]$.

We are to deduce that w belongs to $x_0 + N_B$. Since x_0 is an element of z_0+D, the statement

$$\| Aw - y \| \leqq \| Ax_0 - y \|$$

is a true instance of the foregoing assumption. On the other hand, x_0 belongs to $x_0 + N_B$, whereas w is now in z_0+D. Hence, by the conclusion to the first part of this proof, suitably modified,

$$\forall w \in z_0+D[\ \| Ax_0 - y \| \leqq \| Aw - y \|\ ]$$

holds. For the w in question, it follows that

$$\| Aw - y \| = \| Ax_0 - y \|$$

so that $\|B(x_0 - w)\| = 0$, $B(x_0 - w) = \theta$, $x_0 - w \in N_B$, and w is an element of $x_0 + N_B$.

To prove the last part of the theorem, consider again the true proposition

$$\forall w \in x_0+N_B\ \forall x \in z_0+D[\ \| Ax-y \|^2 = \| B(x-w) \|^2 + \| Aw-y \|^2].$$

If x is an element of z_0+D, then

$$\forall w \in x_0+N_B[\ \| Aw - y \| = \| Ax - y \|\ ]$$

is true if and only if

$$\forall w \in x_0+N_B[\ \| B(x - w) \| = 0]$$

holds. This statement is equivalent to

$$\forall w \in x_0+N_B[x - w \in N_B].$$

Thus, the condition

$$\forall w \in x_0+N_B[\ \| Aw - y \| = \| Ax - y \|\ ]$$

is satisfied if and only if x belongs to $x_0 + N_B$. ∎

When System III has no solution but System I does, a solution of System I minimizes the norm $\| Ax - y \|$ over z_0+D. In this sense, a solution of System I comes as "close" to satisfying System III as does any element of z_0+D. Accordingly, solutions of System I are appropriately termed "best" approximate solutions of System III. Should y happen to be an element of $Az_0 + R_B$, a "best" approximate solution is, in fact, a solution of III.

CHAPTER IV

GENERALIZED INVERSE OPERATORS

At the outset of the preceding chapter we considered a linear operator $\mathcal{A}$ with domain X, and its restriction A to the translation z_0+D of a subset D of X. Other notation was also adopted which we shall continue to adhere to. Recall, in particular, that z_0 denoted some element, and D a linear subset of X. In this chapter an operator $A^{\ddagger}$, called the "generalized inverse" of A, is defined and a number of its properties established.

To facilitate the development, we assume throughout that z_0+D is a subset of $N_B \oplus N_B^{\perp}$. This would be the case, for instance, if N_B were closed and X were complete. For, by Theorem 2.2, $N_B \oplus N_B^{\perp}$ would then be all of X. In any case, let P_N designate the projection operator onto N_B.

<u>THEOREM</u> <u>4.1</u>: For each y in the set $Az_0 + R_B \oplus R_B^{\perp}$, the system:

$$\text{VI} \qquad \begin{cases} Ax - y \in R_B^{\perp} \\ P_N x = \Phi \\ x \in z_0 + D \end{cases}$$

has a unique solution.

PROOF: Since y is an element of $Az_o + R_B \oplus R_B^{\perp}$, y must equal $Az_o + s$ where s is an element of $R_B \oplus R_B^{\perp}$. So the difference $Az_o - y = -s$ is in the set $R_B \oplus R_B^{\perp}$. Consequently, by Theorems 3.2 and 3.6, there is an x_o in $z_o + D$ such that $x_o + N_B$ is the solution set of System I.

Since $z_o + D$ is assumed to be a subset of the domain $N_B \oplus N_B^{\perp}$ of P_N, the projection $P_N x_o$ of x_o onto N_B is defined. Now, the element $x_o - P_N x_o$ belongs to $x_o + N_B$. Thus, the first and last relations of System VI are satisfied by $x_o - P_N x_o$. Since it is also true that

$$\begin{aligned} P_N(x_o - P_N x_o) &= P_N x_o - P_N(P_N x_o) \\ &= P_N x_o - P_N x_o \\ &= \Phi, \end{aligned}$$

$x_o - P_N x_o$ is a solution of System VI.

To show that this solution is unique, we first observe that all solutions of VI must be elements of the set $x_o + N_B$. Hence, they have the form $x_o + u$, where u is an element of N_B. Upon requiring $P_N(x_o + u)$ to equal Φ, we obtain

$$\begin{aligned} \Phi &= P_N(x_o + u) \\ &= P_N x_o + P_N u \\ &= P_N x_o + u. \end{aligned}$$

These equations are satisfied if and only if $u = -P_N x_o$.

Of all the members of $x_0 + N_B$, only $x_0 - P_N x_0$ fulfills each condition of System VI. Therefore, the solution of VI is unique.∎

Having proven that for each y in the set $Az_0 + R_B \oplus R_B^{\perp}$ there exists a unique solution of System VI, we now use this result to define the <u>generalized</u> <u>inverse</u> operator $A^{\ddagger}$ corresponding to the operator A.

<u>DEFINITION</u>: Let $\mathcal{a}$ be a linear operator defined on X. Designate the restriction of $\mathcal{a}$ to z_0+D by A, where z_0 is some point and D a linear subset of X. The generalized inverse $A^{\ddagger}$ of A is an operator having $Az_0 + R_B \oplus R_B^{\perp}$ as domain. At each point of its domain the value of $A^{\ddagger}$ is the unique solution of System VI.

If Y is complete, the domain $Az_0 + R_B \oplus R_B^{\perp}$ of $A^{\ddagger}$ is dense in Y (Theorem A.3). The range of $A^{\ddagger}$ is a subset of z_0+D. The following theorems shed some light on just how $A^{\ddagger}$ differs in behavior from an ordinary inverse operator. They also reveal that $A^{\ddagger}$ is not, in general, a linear operator. This is to be expected since the domain $Az_0 + R_B \oplus R_B^{\perp}$ of $A^{\ddagger}$ is usually nonlinear.

THEOREM 4.2: $A^{\ddagger}(Az_o) = z_o - P_N z_o$.

PROOF: To prove this theorem we shall show that $x_1 = z_o - P_N z_o$ satisfies System VI when y is replaced by Az_o. First, note that

$$\begin{aligned} A(z_o - P_N z_o) - Az_o &= Az_o - B(P_N z_o) - Az_o \\ &= -B(P_N z_o) \\ &= \theta \end{aligned}$$

which is an element of $R_B^{\perp}$. Secondly, observe that

$$P_N(z_o - P_N z_o) = P_N z_o - P_N z_o = \Phi.$$

Finally, $z_o - P_N z_o$ is an element of $z_o + D$ since $P_N z_o$ is in N_B, a subset of D. Therefore, x_1 is the solution of System VI when y equals Az_o. Since this solution is unique, $A^{\ddagger} Az_o = z_o - P_N z_o$. ∎

THEOREM 4.3: Let y_1 and y_2 be arbitrary elements of $Az_o + R_B \oplus R_B^{\perp}$, and α and β arbitrary elements of the field. Then, $\alpha y_1 + \beta y_2 + [1-\alpha-\beta]Az_o$ is an element of $Az_o + R_B \oplus R_B^{\perp}$ and

$$A^{\ddagger}(\alpha y_1 + \beta y_2 + [1-\alpha-\beta]Az_o) = \alpha A^{\ddagger} y_1 + \beta A^{\ddagger} y_2 + [1-\alpha-\beta]A^{\ddagger} Az_o.$$

PROOF: By assumption, y_1 and y_2 are arbitrary elements of $Az_o + R_B \oplus R_B^{\perp}$ and α and β are arbitrary elements of the field. Thus, $\alpha y_1 + \beta y_2 + [1-\alpha-\beta]Az_o$ is an element of $\alpha Az_o + \beta Az_o + [1-\alpha-\beta]Az_o + R_B \oplus R_B^{\perp}$ which is the set $Az_o + R_B \oplus R_B^{\perp}$.

We next show that whenever y in System VI is replaced by $\alpha y_1 + \beta y_2 + [1-\alpha-\beta]Az_o$,

$$x_1 = \alpha A^{\#} y_1 + \beta A^{\#} y_2 + [1-\alpha-\beta]A^{\#} Az_o$$

is the solution of that system.

Since $AA^{\#} y_1 - y_1$ and $AA^{\#} y_2 - y_2$ are elements of $R_B^{\perp}$, and

$$AA^{\#} Az_o = A(z_o - P_N z_o)$$
$$= Az_o$$

by Theorem 4.2,

$$A(\alpha A^{\#} y_1 + \beta A^{\#} y_2 + [1-\alpha-\beta]A^{\#} Az_o) - (\alpha y_1 + \beta y_2 + [1-\alpha-\beta]Az_o)$$
$$= \alpha a A^{\#} y_1 + \beta a A^{\#} y_2 + [1-\alpha-\beta] a A^{\#} Az_o - \alpha y_1 - \beta y_2 - [1-\alpha-\beta]Az_o$$
$$= \alpha AA^{\#} y_1 + \beta AA^{\#} y_2 + [1-\alpha-\beta]AA^{\#} Az_o - \alpha y_1 - \beta y_2 - [1-\alpha-\beta]Az_o$$
$$= \alpha(AA^{\#} y_1 - y_1) + \beta(AA^{\#} y_2 - y_2)$$

is an element of $R_B^{\perp}$. Thus,

$$Ax_1 - (\alpha y_1 + \beta y_2 + [1-\alpha-\beta]Az_o) \in R_B^{\perp}$$

holds.

Now, x_1 satisfies the equation $P_N x = \Phi$, because $P_N A^{\#} y_1$, $P_N A^{\#} y_2$, and $P_N A^{\#} Az_o$ are all equal to Φ, and P_N is a linear operator.

To prove that x_1 is an element of $z_o + D$, observe that

$A^{\ddagger}y_1$, $A^{\ddagger}y_2$, and $A^{\ddagger}Az_0$ are all elements of z_0+D. Consequently,

$$x_1 = \alpha A^{\ddagger}y_1 + \beta A^{\ddagger}y_2 + [1-\alpha-\beta]A^{\ddagger}Az_0$$

is an element of

$$(\alpha z_0 + \beta z_0 + [1-\alpha-\beta]z_0) + D = z_0 + D. \blacksquare$$

The statement of equality in Theorem 4.3 is somewhat deceptive in that the rule of $A^{\ddagger}$ resembles the behavior of that for a linear operator. But, notice that the coefficient of Az_0, in the left hand member, is dependent upon the coefficients of y_1 and y_2. The question of just when $A^{\ddagger}$ reduces to a linear operator is answered by the ensuing theorem.

<u>THEOREM</u> <u>4.4</u>: $A^{\ddagger}$ is a linear operator if and only if z_0 is an element of D.

<u>PROOF</u>: Assume $A^{\ddagger}$ is a linear operator. Then the domain of $A^{\ddagger}$ must be linear; thus $A^{\ddagger}\theta$ is defined. In fact, $A^{\ddagger}\theta = \Phi$; for the rule of $A^{\ddagger}$ is linear too. Therefore, Φ is an element of z_0+D. Consequently, Φ is equal to z_0+d for some d in D. We conclude that $d = -z_0$ belongs to D. Since D is linear, z_0 is also an element of D.

Now, suppose that z_o is an element of D. Our objective is to verify that $A^{\ddagger}$ is a linear operator. To do this, we shall show that both the domain and the rule of operation of $A^{\ddagger}$ are linear. For any z_o in D, $Az_o = Bz_o$ is an element of R_B. Hence, $Az_o + R_B \oplus R_B^{\perp}$ is the same set as $R_B \oplus R_B^{\perp}$. Thus, the domain of $A^{\ddagger}$ is linear. To consummate the proof, we next demonstrate that the rule of operation of $A^{\ddagger}$ is linear.

Let α and β be arbitrary elements of the field, and y_1 and y_2 arbitrary elements of the domain $R_B \oplus R_B^{\perp}$ of $A^{\ddagger}$. By definition, $x_1 = A^{\ddagger}(\alpha y_1 + \beta y_2)$ is the unique solution of System VI when y is replaced by $\alpha y_1 + \beta y_2$. We proceed to prove that $x_2 = \alpha A^{\ddagger} y_1 + \beta A^{\ddagger} y_2$ is also a solution of System VI for the same value of y.

The definitions of $A^{\ddagger} y_1$ and $A^{\ddagger} y_2$ ensure that $AA^{\ddagger} y_1 - y_1$ and $AA^{\ddagger} y_2 - y_2$ are elements of $R_B^{\perp}$. Thus, using the fact that A is now linear because it is the same operator as B, we find that

$$\begin{aligned} Ax_2 - (\alpha y_1 + \beta y_2) &= A(\alpha A^{\ddagger} y_1 + \beta A^{\ddagger} y_2) - (\alpha y_1 + \beta y_2) \\ &= \alpha AA^{\ddagger} y_1 + \beta AA^{\ddagger} y_2 - \alpha y_1 - \beta y_2 \\ &= \alpha (AA^{\ddagger} y_1 - y_1) + \beta (AA^{\ddagger} y_2 - y_2) \end{aligned}$$

is also an element of $R_B^{\perp}$. In addition,

$$P_N x_2 = P_N(\alpha A^{\ddagger} y_1 + \beta A^{\ddagger} y_2)$$
$$= \alpha P_N A^{\ddagger} y_1 + \beta P_N A^{\ddagger} y_2 = \Phi.$$

Inasmuch as $A^{\ddagger} y_1$ and $A^{\ddagger} y_2$ are elements of $z_0 + D = D$, and D is linear, it is evident that $x_2 = \alpha A^{\ddagger} y_1 + \beta A^{\ddagger} y_2$ is in $z_0 + D$. Thus, x_2 and x_1 are solutions of System VI for the same value of y. It follows that

$$A^{\ddagger}(\alpha y_1 + \beta y_2) = \alpha A^{\ddagger} y_1 + \beta A^{\ddagger} y_2 ;$$

that is, the rule of operation of $A^{\ddagger}$ is linear.∎

Many of the most important properties of the operator $A^{\ddagger}$ are specified by the following theorems.

<u>THEOREM</u> <u>4.5</u>: For every x in $z_0 + D$,

$$A^{\ddagger} A x = x - P_N x.$$

<u>PROOF</u>: Let x be an element of $z_0 + D$. Then, Ax is an element of $Az_0 + R_B$, so $A^{\ddagger} A x$ is defined. We shall prove that $x - P_N x$ is a solution of System VI when y is replaced by Ax. First of all,

$$A(x - P_N x) - Ax = -BP_N x$$
$$= \theta$$

is an element of $R_B^{\perp}$. What is more, the equalities

$$P_N(x - P_N x) = P_N x - P_N x$$
$$= \Phi$$

holds. Finally, $x - P_N x$ is an element of z_0+D because x is in z_0+D and $P_N x$ is in D. Thus, $A^{\ddagger}Ax = x - P_N x$ is verified.▮

<u>THEOREM 4.6</u>: For every x in z_0+D,

$$A^{\ddagger}Ax = P_N^{\perp}x$$

where $P_N^{\perp}x$ is the projection operator onto $N_B^{\perp}$.

<u>PROOF</u>: Given any x in z_0+D, we know from Theorem 4.5 that $A^{\ddagger}Ax = x - P_N x$. But, keeping in mind that z_0+D is a subset of $N_B \oplus N_B^{\perp}$,

$$x = (x - P_N x) + P_N x$$

where $x - P_N x$ is an element of $N_B^{\perp}$ and $P_N x$ is in N_B. By Theorem A.2 (see the proof of (iii) in particular), $\overline{N_B}$ is a subset of $(N_B^{\perp})^{\perp}$. Thus, $P_N x$ is in $(N_B^{\perp})^{\perp}$. Therefore, by the definition of the projection operator, $x - P_N x = P_N^{\perp}x$ holds for every x in z_0+D.▮

<u>THEOREM 4.7</u>: Let x be an element of z_0+D. Then,

$$AA^{\ddagger}Ax = Ax.$$

PROOF: According to Theorem 4.5, $A^{\#}Ax = x - P_N x$. Thus, for every x in $z_o + D$, we have

$$AA^{\#}Ax = a(x - P_N x)$$
$$= Ax - BP_N x$$
$$= Ax.$$

THEOREM 4.8: For every y in $Az_o + R_B \oplus R_B^{\perp}$,

$$AA^{\#}y = y - P_R^{\perp}(y - Az_o)$$

where $P_R^{\perp}$ is the projection operator onto $R_B^{\perp}$.

PROOF: Note that $A^{\#}y$ is an element of $z_o + D$ so that $AA^{\#}y$ is defined. We know that $AA^{\#}y - y = v$, for some element v of $R_B^{\perp}$, from the definition of $A^{\#}y$. With y fixed, the corresponding v may be determined as follows. Make use of the preceding relation to express $y - Az_o$ as

$$y - Az_o = y + v - Az_o - v$$
$$= AA^{\#}y - Az_o - v.$$

Since $A^{\#}y$ is an element of $z_o + D$, the equality $AA^{\#}y = Az_o + Bd$ is valid for some element d of D. Thus,

$$y - Az_o = Bd - v$$

where Bd is in R_B and, as has been said, v is an element of $R_B^{\perp}$. Since R_B and $R_B^{\perp}$ are subsets of $R_B^{\perp} \oplus (R_B^{\perp})^{\perp}$, Bd, v, and perforce $y - Az_o$ are in the domain of $P_R^{\perp}$. Consequently,

$$P_R^{\perp}(y - Az_o) = P_R^{\perp}(Bd - v)$$
$$= P_R^{\perp}Bd - P_R^{\perp}v$$
$$= -v.$$

Therefore, $v = -P_R^{\perp}(y - Az_o)$ and

$$y - Az_o = AA^{\ddagger}y - Az_o + P_R^{\perp}(y - Az_o).$$

Solving for $AA^{\ddagger}y$ we have $AA^{\ddagger}y = y - P_R^{\perp}(y - Az_o)$.∎

<u>THEOREM</u> <u>4.9</u>: For every y in $Az_o + R_B \oplus R_B^{\perp}$,

$$AA^{\ddagger}y = Az_o + P_R(y - Az_o)$$

where P_R is the projection operator onto R_B.

<u>PROOF</u>: For any y in $Az_o + R_B \oplus R_B^{\perp}$, the element $y-Az_o$ belongs to $R_B \oplus R_B^{\perp}$. We may therefore employ the definition of the projection operator, and Theorem 4.8, to obtain two equivalent representations for $y-Az_o$; namely,

$$y - Az_o = P_R(y - Az_o) + P_R^{\perp}(y - Az_o)$$
$$= AA^{\ddagger}y - Az_o + P_R^{\perp}(y - Az_o).$$

Upon equating right hand members, and solving for $AA^{\ddagger}y$, we arrive at the conclusion of the theorem.∎

<u>THEOREM</u> <u>4.10</u>: For every y in $Az_o + R_B \oplus R_B^{\perp}$,

$$A^{\ddagger}AA^{\ddagger}y = A^{\ddagger}y.$$

PROOF: Let y_1 be an arbitrary element of $Az_0 + R_B \oplus R_B^\perp$. Then, $AA^\ddagger y_1$ is in $Az_0 + R_B$ and $A^\ddagger AA^\ddagger y_1$ is defined. When y in System VI takes on the value $AA^\ddagger y_1$, $x_1 = A^\ddagger AA^\ddagger y_1$ is the solution of that system by definition. We shall prove that

$$A^\ddagger AA^\ddagger y_1 = A^\ddagger y_1$$

by showing that $x_2 = A^\ddagger y_1$ is also the solution of VI when y has the same value. To do this, we observe that

$$\begin{aligned} Ax_2 - AA^\ddagger y_1 &= AA^\ddagger y_1 - AA^\ddagger y_1 \\ &= \theta \end{aligned}$$

is an element of $R_B^\perp$. From its definition, $A^\ddagger y_1$ is an element of z_0+D, and $P_N A^\ddagger y_1$ equals Φ. Thus, x_2 is the solution x_1 of System VI. Since y_1 was arbitrary, we conclude that for every y in $Az_0 + R_B \oplus R_B^\perp$

$$A^\ddagger AA^\ddagger y = A^\ddagger y$$

is true. ▮

Having now established several properties of $A^\ddagger$, let us analyze how our definition of $A^\ddagger$ generalizes the usual definition of the inverse operator.

The operator A, introduced in Chapter III, has domain z_0+D and range $Az_0 + R_B$. Such an operator has an ordinary inverse A^{-1} if and only if a single element of

z_0+D is mapped onto each element of Az_0+R_B. Such a mapping is said to be one-to-one. When A is one-to-one so is the operator A^{-1}; it maps Az_0+R_B onto z_0+D.

To put matters more concisely, A^{-1} exists if and only if both of the statements

$$\forall x \in z_0+D[A^{-1}Ax = x]$$

and

$$\forall y \in Az_0+R_B[AA^{-1}y = y]$$

are true.

If N_B does not equal $\{\Phi\}$, the first of these statements is false. To see this, let u denote any nonzero element of N_B. Then z_0+u, as well as z_0, belongs to z_0+D. Suppose $A^{-1}Az_0 = z_0$ holds. Using the fact that

$$\begin{aligned} A(z_0+u) &= Az_0+Bu \\ &= Az_0+\theta \\ &= Az_0, \end{aligned}$$

we find that

$$\begin{aligned} A^{-1}A(z_0+u) &= A^{-1}Az_0 \\ &= z_0; \end{aligned}$$

consequently, $A^{-1}A(z_0+u)$ does not equal z_0+u. This proves that A^{-1} exists only if $N_B = \{\Phi\}$.

The converse of this conclusion is also true; that is, A^{-1} exists if $N_B = \{\Phi\}$. For if N_B equals $\{\Phi\}$, and $Ax_1 = Ax_2$ for some elements x_1 and x_2 of z_0+D, then the equalities

$$\theta = Ax_1 - Ax_2$$
$$= B(x_1 - x_2).$$

hold. Consequently, $x_1 - x_2$ is in N_B. That is, $x_1 - x_2$ equals Φ and $x_1 = x_2$. Therefore, A is one-to-one and A^{-1} exists.

In case N_B does not equal $\{\Phi\}$, it is possible to define a generalized inverse $A^{\ddagger}$ of A as a map from $Az_0 + R_B$ into z_0+D for which the statements

$$\forall x \in z_0+D[A^{\ddagger}Ax = x - P_N x]$$

and

$$\forall y \in Az_0+R_B[AA^{\ddagger}y = y]$$

are required to hold.

Actually, we have gone even farther than this. Our generalized inverse operator $A^{\ddagger}$ has $Az_0 + R_B \oplus R_B^{\perp}$ as domain. This set includes $Az_0 + R_B$ as a subset. Indeed, Theorems 4.5 and 4.8 reveal that $A^{\ddagger}$ is a map from $Az_0 + R_B \oplus R_B^{\perp}$ into z_0+D such that both of the statements

$$\forall x \in z_0+D[A^{\ddagger}Ax = x - P_N x]$$

and

$$\forall y \in Az_0+R_B \oplus R_B^{\perp}[AA^{\ddagger}y = y - P_R^{\perp}(y - Az_0)]$$

are valid.

If $N_B = \{\Phi\}$, then $P_N x = \Phi$ for every x in z_0+D. Hence, the first of the preceding pair of statements becomes