ILLINOIS JOURNAL OF MATHEMATICS
Volume 34, No. 3, Fall 1990

SOME OPERATOR INEQUALITIES CONCERNING GENERALIZED INVERSES

BY

P.J. MAHER

In this paper we extend to the von Neumann-Schatten classes some inequalities concerning generalized inverses originally established by Penrose [7] for matrices.

It is well known that if T is any matrix then there exists a matrix T^- such that $TT^-T = T$. The matrix T^- is called a generalized (or, sometimes, a pseudo or inner) inverse of T. A non-invertible matrix has an infinity of generalized inverses, but an invertible matrix has a unique generalized inverse which coincides with its ordinary inverse.

Although the matrix T may have an infinity of generalized inverses, there exists a canonical generalized inverse, called the Moore-Penrose inverse and denoted by T^+, which is uniquely determined by T. Further, as Penrose showed in [7], the Moore-Penrose inverse satisfies the following inequalities: for all X,

$$\|AX - C\|_2 \geq \|AA^+C - C\|_2 \tag{P.1}$$

with equality occurring in (P.1) if and only if $X = A^+C + (I - A^+A)L$ where L is arbitrary; and

$$\|A^+C + (I - A^+A)L\|_2 \geq \|A^+C\|_2 \tag{P.2}$$

with equality occurring in (P.2) if and only if $(I - A^+A)L = 0$. (The only restrictions on the matrices occurring in (P.1) and (P.2) is that they be conformable for multiplication. In (P.1) and (P.2) $\|\cdot\|_2$ denotes the Euclidean norm on matrices.)

This paper contains, in Theorems 2.1 and 2.2, infinite-dimensional extensions of the inequalities (P.1) and (P.2) to the supremum norm on $\mathscr{L}(H)$ and to the von Neumann-Schatten norms $\|\cdot\|_p$, where $2 \leq p < \infty$. The proofs of Theorems 2.1 and 2.2 are an extension of Penrose's original proof of (P.1) and (P.2) and depend on an inequality, viz. Theorem 1.7, about operators having orthogonal ranges/co-ranges.

Received February 29, 1988.

In addition, we obtain local results, using the Aiken, Erdos and Goldstein formula for the derivative of the map $X \mapsto \|X\|_p^p$, where $1 < p < \infty$ [1, Theorem 2.1]. The critical points of the map $X \mapsto \|AX - C\|_p^p$ are classified for $1 < p < \infty$ in Theorem 2.3. Theorem 2.3 shows that for $2 \leq p < \infty$ critical points and global minima coincide. It also extends the global inequality (P.1) to the $1 < p < 2$ case provided the underlying space is finite-dimensional. Theorem 2.3 would appear to be useful for justifying numerical methods of finding generalized inverses.

There are, of course, similar results about minimizing $\|XB - D\|$ and $\|XB - D\|_p$ (see (2.4) and (2.5)).

Section 3 concerns a different problem, that of minimizing $\|T^-\|$ ($\|T^-\|_p$) where T^- ranges over all the generalized inverses (in $\mathscr{C}_p$) of some fixed operator T. Almost inevitably, $\|T^-\|$ ($\|T^-\|_p$) attain their minimums when T^- is the Moore-Penrose inverse T^+ (see Theorem 3.1). Further, $T^- = T^+$ is the only critical point of the map $T^- \mapsto \|T^-\|_p^p$, where $1 < p < \infty$ (see Theorem 3.2). The results of §3 are really finite-dimensional: as is pointed out in §3, if T has a compact generalized inverse then T and T^* must be of finite rank.

This paper originated as part of my Ph.D. thesis. I should like to thank my supervisor, Dr. J.A. Erdos, for the help and encouragement he has so freely and generously given.

After writing this paper I learned of Engl and Nashed's work [4]. Some of their results are similar to the ones obtained here. Thus, they prove a special case (viz. when $p = 2$) of Theorem 3.1 (b) [4, theorem (6.2)]. Their methods are quite different to the ones used in this paper.

1. Preliminaries

An operator T^- is said to be a *generalized inverse* of the operator T if $TT^-T = T$. An operator T in $\mathscr{L}(H)$ has a generalized inverse if and only if $\operatorname{Ran} T$ is closed [9, p. 251, Theorem 12.9]. (In this paper the range of T, denoted $\operatorname{Ran} T$, is the set $\{Tf \colon f \in H\}$. The space of all bounded linear operators on the complex separable Hilbert space H is denoted by $\mathscr{L}(H)$.)

Generalized inverses have applications to operator equations. Penrose showed [6, Theorem 2] that if the operator equation $AXB = C$ has a particular solution, say X_1, and if A and B have generalized inverses, A^- and B^-, then every solution of this equation is given by

$$X = X_1 + L - A^-ALBB^-, \quad L \text{ arbitarary in } \mathscr{L}(H). \tag{1.1}$$

For an operator T in $\mathscr{L}(H)$ with closed range the Moore-Penrose inverse is constructed as follows. Let T_0 be the restriction of T to $(\operatorname{Ker} T)^\perp$. Then T_0 is a bijection from $(\operatorname{Ker} T)^\perp$ onto $\operatorname{Ran} T$. Since $\operatorname{Ran} T$ is closed, T_0 has a bounded inverse T_0^{-1}. Let Q be the projection onto Ran T. The operator T^+, defined by

$$T^+ = T_0^{-1}Q, \tag{1.2}$$

is called the *Moore-Penrose inverse* of T. Clearly, T^+ is a generalized inverse of T. Indeed, T^+ satisfies

$$\begin{aligned} TT^+T &= T && \text{(i)} \\ T^+TT^+ &= T^+ && \text{(ii)} \\ (TT^+)^* &= TT^+ && \text{(iii)} \\ (T^+T)^* &= T^+T && \text{(iv)} \end{aligned} \tag{1.3}$$

and, further, the operator T^+ is uniquely determined by the properties (i), (ii), (iii), (iv) [6, Theorem 1].

Often we shall refer to generalized inverses that satisfy some of the properties (i),...,(iv) of (1.3). If an operator T^- satisfies properties (i) and (ii) of (1.3) we shall say that T^- is a (i), (ii) inverse of T; if T^- satisfies (i) and (iii) we shall say that T^- is a (i), (iii) inverse of T, etc.

Lemma 1.4 (required in §3) relates an arbitrary generalized inverse T^- of an operator T, having closed range, to its Moore-Penrose inverse T^+. Although Lemma 1.4 is an easy enough consequence of (1.2), a proof of it is given as it does not seem to figure in the literature. It says that T^+ is the same as T^- acting on Ran T and then projected onto $(\text{Ker } T)^\perp$.

LEMMA 1.4. *Let T in $\mathscr{L}(H)$ have closed range, let T^- be a generalized inverse of T and T^+ be its Moore-Penrose inverse and let P be the projection onto* Ker T *and Q be the projection onto* Ran T. *Then*:

(a) $T^+ = (I - P)T^-Q$;

(b) $T^- = T^+$ *if and only if* $T^-(I - Q) = 0$ *and* $PT^- = 0$;

(c) $\{T^-: TT^-T = T\} = \{T^+ + S: S \in \mathscr{L}(H)$ *and* $(I - P)SQ = 0\}$;

(d) *If T^- and $\tilde{T}^-$ are any two generalized inverses of T then $T^- - \tilde{T}^- = S$ where $(I - P)SQ = 0$.*

Proof. (a) For arbitrary g in H, let $g = Tf + h$ where $f \in (\text{Ker } T)^\perp$ and $h \in (\text{Ran } T)^\perp$ so that $Qg = Tf$. By (1.2), $T^+g = f$. On the other hand,

$$(I - P)T^-Qg = (I - P)T^-Tf = (I - P)(v + w) = v$$

where $T^-Tf = v + w$, $v \in (\text{Ker } T)^\perp$ and $w \in \text{Ker } T$; and $v = f$ as desired since $v - f \in (\text{Ker } T)^\perp$ and $v - f \in \text{Ker } T$ (for $Tv = T(v + w) = T(T^-Tf) = Tf$).

(b) This follows immediately from (a).

(c) If T^- is a generalized inverse of T, let $S = T^- - T^+$; then by (a), $(I - P)SQ = 0$. Conversely, if S satisfies $(I - P)SQ = 0$ then $T^+ + S$ is a generalized inverse of T since $T(T^+ + S)T = TT^+T = T$ (for, since $T = T(I - P) = QT$, we have $TST = T(I - P)SQT = 0$).

(d) This follows immediately from (c). ■

For details of the von Neumann-Schatten classes $\mathscr{C}_p$ and the norms $\|\cdot\|_p$ see [8, Chapter 2], [3, Chapter XI], or [5, §2].

The local results of this paper hinge on the Aiken, Erdos and Goldstein differentiation result which we now state. (The real part of a complex number z will be denoted by $\mathscr{R}z$.)

THEOREM 1.5 [1, Theorem 2.1]. *If $1 < p < \infty$, the map $X \mapsto \|X\|_p^p$ (from $\mathscr{C}_p$ to $\mathbf{R}^+$) is differentiable with derivative D_X at X given by*

$$D_X(T) = p\mathscr{R}\tau\left[|X|^{p-1}U^*T\right]$$

where $\tau[\ ]$ denotes trace and where $X = U|X|$ is the polar decomposition of X. If the underlying Hilbert space is finite-dimensional, the same result holds for $0 < p \leq 1$ at every invertible element X.

Next, we come to the inequality about operators in $\mathscr{L}(H)$ with orthogonal ranges/co-ranges. Observe that if Ran $A \perp$ Ran B and Ran $A^* \perp$ Ran B^* then the space H can be decomposed so that either

$$A + B = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \quad \text{or} \quad A + B = \begin{bmatrix} 0 & B \\ A & 0 \end{bmatrix}. \tag{1.6}$$

THEOREM 1.7 (a) *If* Ran $A \perp$ Ran B (*or, if* Ran $A^* \perp$ Ran B^*) *then*

$$\|A + B\| \geq \max\{\|A\|, \|B\|\}; \tag{1}$$

(b) *and if both* Ran $A \perp$ Ran B *and* Ran $A^* \perp$ Ran B^* *then equality holds in* (1).

(c) *If $A + B \in \mathscr{C}_p$ and* Ran $A \perp$ Ran B (*or, if* Ran $A^* \perp$ Ran B^*) *then $A \in \mathscr{C}_p$ and $B \in \mathscr{C}_p$ where $0 < p < \infty$, and*

$$\|A + B\|_p^p \geq \|A\|_p^p + \|B\|_p^p \tag{2}$$

for $2 \leqq p < \infty$, with equality in the $p = 2$ case;

(d) *and if both* Ran $A \perp$ Ran B *and* Ran $A^* \perp$ Ran B^* *then equality holds in* (2) *for $0 < p < \infty$.*

Proof. (a) This is obvious.

(b) and (d) The equalities here follow from considering the matricial representation (1.6) of $A + B$.

(c) If $A + B \in \mathscr{C}_p$ for $0 < p < \infty$ and E is the projection onto Ran A then $A = E(A + B) \in \mathscr{C}_p$. Hence $B \in \mathscr{C}_p$.

To prove the inequality (2) we require the following result from [2, Lemma 4]: if $T \in \mathscr{C}_p$ where $2 \leq p < \infty$ and if $\{\phi_k\}$ is an arbitrary orthonormal basis of H then

$$\|T\|_p^p \geq \sum_k \|T\phi_k\|^p \tag{3}$$

with equality in the $p = 2$ case.

Suppose Ran A^* ($= (\operatorname{Ker} A)^\perp$) $\perp$ Ran B^* ($= (\operatorname{Ker} B)^\perp$). It follows from the Schmidt expansion that there exist orthonormal bases $\{x_i\}$, $\{y_j\}$ of $(\operatorname{Ker} A)^\perp$ and $(\operatorname{Ker} B)^\perp$ respectively such that

$$\|A\|_p^p = \sum_i \|Ax_i\|^p \quad \text{and} \quad \|B\|_p^p = \sum_j \|By_j\|^p.$$

Since $\{x_i\} \cup \{y_j\}$ is orthonormal, (2) follows from (3) (let $\{\phi_k\} \supseteq \{x_i\} \cup \{y_j\}$). If, instead, Ran $A \perp$ Ran B take adjoints. ■

The next example shows that if only Ran $A \perp$ Ran B and Ran $A^* \not\perp$ Ran B^* then the inequality (2) of Theorem (1.7) (c) may not hold if $p < 2$. Let $H = \mathbf{C}^2$,

$$A = \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 \\ b & 0 \end{bmatrix}$$

where $a \neq 0$ and $b \neq 0$. For $p < 2$, from Jensen's inequality [9, p. 3] we get $\|A + B\|_p^p \leq \|A\|_p^p + \|B\|_p^p$: for an instance when this inequality is strict take $a = 3$, $b = 4$ and $p = \frac{3}{2}$.

2. On minimizing $\|AX - C\|_p$

We now extend Penrose's inequalities (P.1) and (P.2) to the supremum norm and to the $\mathscr{C}_p$ norm.

THEOREM 2.1. *Let A have closed range and have a* (i), (iii) *inverse A^-. Then for all X in $\mathscr{L}(H)$,*

$$\|AX - C\| \geq \|AA^-C - C\| \tag{1}$$

with equality occurring in (1) *if $X = A^-C + (I - A^-A)L$ where L is arbitrary*;

and on taking A^- as A^+,

$$\|A^+C + (I - A^+A)L\| \geq \|A^+C\|. \tag{2}$$

Proof. Consider the identity

$$AX - C = (AX - AA^-C) + (AA^-C - C).$$

Here, $\operatorname{Ran}(AX - AA^-C) \perp \operatorname{Ran}(AA^-C - C)$ because, since A has A^- as a (i), (iii) inverse,

$$(AA^-C - C)^*(AX - AA^-C) = C^*(AA^- - I)A(X - A^-C) = 0.$$

The inequality (1) now follows from Theorem 1.7 (a). If X is given by $X = A^-C + (I - A^-A)L$ then, since A^- is a (i) inverse, $AX = AA^-C$; hence, equality occurs in (1).

If $A^- = A^+$ then $\operatorname{Ran} A^+C \perp \operatorname{Ran}(I - A^+A)L$ (since A^+ is, in particular, a (iii), (iv) inverse of A). The inequality (2) now comes from Theorem 1.7 (a). ∎

THEOREM 2.2. *Let A have closed range and have a* (i), (iii) *inverse A^- and let X be such that $AX - C \in \mathscr{C}_p$. Then $AA^-C - C \in \mathscr{C}_p$ where $0 < p < \infty$; and*

$$\|AX - C\|_p \geq \|AA^-C - C\|_p \tag{1}$$

for $2 \leq p < \infty$, with equality occurring in (1) *if and only if*

$$X = A^-C + (I - A^-A)L$$

where L is arbitrary. If A^- is taken as A^+ and if $A^+C + (I - A^+A)L \in \mathscr{C}_p$ then $A^+C \in \mathscr{C}_p$ where $0 < p < \infty$; and

$$\|A^+C + (I - A^+A)L\|_p \geq \|A^+C\|_p \tag{2}$$

for $2 \leq p < \infty$, with equality occurring in (2) *if and only if $(I - A^+A)L = 0$.*

Proof. If $AX - C \in \mathscr{C}_p$ then, as $\operatorname{Ran}(AX - AA^-C) \perp \operatorname{Ran}(AA^-C - C)$ in the identity $AX - C = (AX - AA^-C) + (AA^-C - C)$, it follows from Theorem 1.7 (c) that $AX - AA^-C \in \mathscr{C}_p$, $AA^-C - C \in \mathscr{C}_p$ where $0 < p < \infty$; and

$$\|AX - C\|_p^p \geq \|AX - AA^-C\|_p^p + \|AA^-C - C\|_p^p \geq \|AA^-C - C\|_p^p \tag{3}$$

for $2 \leq p < \infty$. This gives (1). If equality occurs in (1) then equality occurs throughout (3). So, $AX = AA^-C$; so, by (1.1), $X = A^-C + (I - A^-A)L$.

The proof of (2), and the uniqueness statement there, is similar. ■

Theorems 2.1 and 2.2 both have a simple geometric interpretation. Denote the supremum norm $\|\cdot\|$ on $\mathscr{L}(H)$ by $\|\cdot\|_\infty$. Then Theorems 2.1 and 2.2 say that the quantity $\|AX - C\|_p$, where $2 \le p \le \infty$, is minimized when (and, for $2 \le p < \infty$, only when)

$$X = A^- C + M$$

where $M (= (I - A^- A)L)$ is any operator such that $\operatorname{Ran} M \subseteq \operatorname{Ker} A$.

Next we give the local result corresponding to Theorem 2.2.

THEOREM 2.3. *Let A have closed range and have a* (i), (iii) *inverse A^-. Let X vary such that $AX - C \in \mathscr{C}_p$, where $1 < p < \infty$, and let*

$$F_p \colon X \mapsto \|AX - C\|_p^p.$$

Then:

(a) *for $1 < p < \infty$, V is a critical point of F_p if and only if*

$$V = A^- C + (I - A^- A) L$$

where L is arbitrary in $\mathscr{L}(H)$;

(b) *for $2 \le p < \infty$, V is a global minimum of F_p if and only if V is a critical point of F_p;*

(c) *for $1 < p < 2$, the same result as in* (b) *holds provided the underlying space is finite-dimensional.*

Proof. (a) Let V be a critical point of F_p. Let S be an arbitrary increment of V, that is, S is arbitrary provided $A(V + S) - C \in \mathscr{C}_p$, equivalently, $AS \in \mathscr{C}_p$. From Theorem 1.5 we get

$$0 = D_{AV-C}(AS) = p\mathscr{R}\tau\left[|AV - C|^{p-1} U^*(AS)\right] \tag{1}$$

where $AV - C = U|AV - C|$ is the polar decomposition of $AV - C$. Take $S = \lambda(f \otimes g)$ where $\lambda \in \mathbf{C}$ and f and g are arbitrary vectors. (The operator $x \mapsto \langle x, f\rangle g$ is denoted by $f \otimes g$. Note that $\tau[T(f \otimes g)] = \langle Tg, f\rangle$; cf. [8, pp. 73, 90].) From (1) we get

$$0 = \mathscr{R}\lambda\left\langle |AV - C|^{p-1} U^* A g, f\right\rangle.$$

Since λ, f and g are arbitrary,

$$|AV - C|^{p-1} U^* A = 0.$$

As $\operatorname{Ker}|AV - C|^{p-1} = \operatorname{Ker}|AV - C|$, we have $|AV - C|U^*A = 0$, that is, $A^*(AV - C) = 0$. This equation is the same as $AV = AA^-C$ (for since A^- is a (i), (iii) inverse of A, $(A^-)^*A^*A = A$). By (1.1), $V = A^-C + (I - A^-A)L$.
Conversely, if $V = A^-C + (I - A^-A)L$ then

$$0 = |AV - C|U^*A = |AV - C|^{p-1}U^*A.$$

Hence (cf. (1)), $D_{AV-C}(AS) = 0$ for all S. Thus, V is a critical point of F_p.
(b) This follows from (a) and from Theorem 2.2.
(c) We show that F_p: $X \mapsto \|AX - C\|_p^p$ attains its infimum. Once this is proved the assertion in (c) will follow immediately from (a) since a global minimum is a critical point. Let $\dim H = n < \infty$ and $\alpha = \inf\|AX - C\|_p$. Then there exists a sequence $\{X_i\}$ such that $\|AX_i - C\|_p \to \alpha$. Define the sequence $\{Y_i\}$ thus: $Y_i = A^-AX_i$. Then $AY_i = AX_i$ so that $\|AY_i - C\|_p \to \alpha$. Further, $\{Y_i\}$ is norm-bounded: for i sufficiently large,

$$\|Y_i\| \le \|A^-\| \, \|AX_i\| \le \|A^-\|(\alpha + \varepsilon + \|C\|)$$

(since $\|AX_i\| \le \|AX_i - C\|_p + \|C\|$; recall that $\|\cdot\| \le \|\cdot\|_p$). The sequence $\{\|Y_i\|\}$ is thus bounded where $Y_i \in \mathbf{C}^{n^2}$ $(\simeq \mathscr{L}(H))$. Hence $\{Y_i\}$ contains a subsequence $\{Y_{i_j}\}$ such that $Y_{i_j} \to Y$ where $\|AY - C\|_p = \alpha$. Thus, F_p attains its infimum. ∎

There is no $0 < p \le 1$ version of Theorem 2.3 since the condition that $AV - C$ be invertible, where V is a critical point of F_p, leads to the triviality that $A = 0$ (for if V is a critical point of F_p then, as in Theorem 2.3 (a), $A^*(AV - C) = 0$).
Naturally, there are similar results about $XB - D$ with similar proofs. We now state the global result (2.4). If B has closed range and has a (i), (iv) inverse B^- then

$$\|XB - D\|_p \ge \|DB^-B - D\|_p \tag{2.4}$$

for $2 \le p \le \infty$, with equality occurring in (2.4) if

$$X = DB^- + L(I - BB^-)$$

where L is arbitrary; and when B^- is taken as B^+ then DB^+ is a least such global minimizer in $\|\cdot\|_p$. Thus, the quantity $\|XB - D\|_p$, $2 \le p \le \infty$, is minimized when $X = DB + N$ where N is any operator such that Ker $N \supseteq \operatorname{Ran} B$. (If $2 \le p < \infty$, we assume $XB - D \in C_p$. In that case there are the obvious uniqueness assertions.)
We next state the analogous local result: for $1 < p < \infty$,

(2.5) V is a critical point of $X \mapsto \|XB - D\|_p^p$ if and only if $V = DB^- + L(I - BB^-)$, L arbitrary.

Hence (cf. Theorem 2.3 (c)) the inequality (2.4) also holds for $1 < p < \infty$ in finite-dimensions.

As for minimizing $\|AXB - D\|_p$, Penrose [7, cf. Corollary 1] obtained the following $p = 2$ result for matrices: if A^- is a (i), (iii), inverse of A and B^- is a (i), (iv) inverse of B then for all X,

$$\|AXB - C\|_2 \geq \|AA^-CB^-B - C\|_2 \tag{P.3}$$

with equality occurring in (P.3) if and only if $X = A^-CB^- + L - A^-ALBB^-$ where L is arbitrary; and if $A^- = A^+$ and $B^- = B^+$ then A^+CB^+ is the least such global minimizer in $\|\cdot\|_2$. Obviously, (P.3) goes through to $\mathscr{C}_2$. Thus, $\|AXB - D\|_2$ is minimized when and only when $X = A^-CB^- + K$ where K is any operator such that $\operatorname{Ran}(KB) \subseteq \operatorname{Ker} A$ and $\operatorname{Ran} B \subseteq \operatorname{Ker}(AK)$. This result has the following local variant.

THEOREM 2.6. *Let A and B have closed ranges and let A^- be a* (i), (iii) *inverse of A and B^- be a* (i), (iv) *inverse of B. Let X vary such that $AXB - C \in \mathscr{C}_2$. Then the map $X \mapsto \|AXB - C\|_2^2$ has a critical point at V if and only if $V = A^-CB^- + L - A^-ALBB^-$ where L is arbitrary.*

Proof. We sketch the argument for arbitrary p, where $1 < p < \infty$. It follows, as in Theorem 2.3 (a), that V is a critical point of

$$X \mapsto \|AXB - C\|_p^p$$

if and only if

$$B|AVB - C|^{p-1}U^*A = 0$$

(where $AVB - C = U|AVB - C|$ is the polar decomposition of $AVB - C$). Hence, if $p = 2$, $A^*AVBB^* = A^*CB^*$, that is, $AVB = AA^-CB^-B$. By (1.1), $V = A^-CB^- + L - A^-ALBB^-$. ∎

Do (P.3) and Theorem 2.6 hold for arbitrary p?

3. On minimizing $\|T^-\|_p$

THEOREM 3.1. *Let T be fixed in $\mathscr{L}(H)$ with closed range and with Moore-Penrose inverse T^+. Then*:

(a) *for every generalized inverse T^- of T*,

$$\|T^+\| \leq \|T^-\|;$$

(b) *and if* $T^- \in \mathscr{C}_p$, *where* $0 < p < \infty$, *then* $T^+ \in \mathscr{C}_p$ *and*

$$\|T^+\|_p \leq \|T^-\|_p \tag{1}$$

with, for $1 < p < \infty$, *equality occurring in* (1) *if and only if* $T^- = T^+$.

Proof. (a) Let P be the projection onto Ker T and Q be the projection onto Ran T. Then by Lemma 1.4 (a), $T^+ = (I - P)T^-Q$ and so

$$\|T^+\| \leq \|I - P\| \, \|T^-\| \, \|Q\| \leq \|T^-\|.$$

(b) If $T^- \in \mathscr{C}_p$ then $T^+ = (I - P)T^-Q \in \mathscr{C}_p$. Now,

$$\|ABC\|_p \leq \|A\| \, \|B\|_p \|C\|$$

for any operator B in $\mathscr{C}_p$ where $0 < p < \infty$ [3, p. 1093]. Hence

$$\|T^+\|_p \leq \|I - P\| \, \|T^-\|_p \|Q\| \leq \|T^-\|_p.$$

For $1 < p < \infty$, there is at most one global minimizer since the set of generalized inverses of T in $\mathscr{C}_p$ is convex (for if $\tilde{T}^-$ and $\hat{T}^-$ are two such generalized inverses of T, so is $\alpha\tilde{T}^- + (1 - \alpha)\hat{T}^-$. ∎

Theorem 3.2. *Let* T *be fixed in* $\mathscr{L}(H)$ *with closed range and with Moore-Penrose inverse* T^+. *Then if* T^- *varies over those generalized inverses of* T *in* $\mathscr{C}_p$, *where* $1 < p < \infty$, *the map* $G_p \colon T^- \mapsto \|T^-\|_p^p$ *has a critical point at* $\tilde{T}^-$ *if and only if* $\tilde{T}^- = T^+$.

Proof. By Theorem 3.1 (b), T^+ is a global minimizer and hence, for $1 < p < \infty$, a critical point of G_p.

Conversely, let $\tilde{T}^-$ be a critical point of G_p. Let S be an arbitrary increment of $\tilde{T}^-$. Thus, S is arbitrary subject to the condition that $\tilde{T}^- + S$ is also a generalized inverse of T in $\mathscr{C}_p$. Hence, $S \in \mathscr{C}_p$ and, by Lemma 1.4 (d), we have $(I - P)SQ = 0$ where P and Q are the projections onto Ker T and Ran T respectively. Take $S = f \otimes g$ where f and g are arbitrary vectors such that

$$0 = (I - P)(f \otimes g)Q.$$

Thus, $0 = (Qf) \otimes ((I - P)g)$. Hence, $f \in \text{Ran}(I - Q)$ or $g \in \text{Ran}\, P$. Now, with f in $\text{Ran}(I - Q)$ then g can be arbitrary; and with g in Ran P then f can be arbitrary. Thus, we may assume that $f \in \text{Ran}(I - Q)$ and $g \in \text{Ran}\, P$.

As $\tilde{T}^-$ is a critical point of G_p, it follows from Theorem 1.5 that

$$0 = p\mathscr{R}\tau\left[|\tilde{T}^-|^{p-1}U^*(S)\right] = p\mathscr{R}\langle|\tilde{T}^-|^{p-1}U^*g, f\rangle \qquad (1)$$

where $\tilde{T}^- = U|\tilde{T}^-|$ is the polar decomposition of $\tilde{T}^-$ (hence $\operatorname{Ker} U = \operatorname{Ker}|\tilde{T}^-|$). To prove $\tilde{T}^- = T^g$ we shall appeal to Lemma 1.4 (b).

Since $f \in \operatorname{Ran}(I - Q)$ then $f = (I - Q)y$ for arbitrary y. Substituting this expression for f into (1) we get

$$\mathscr{R}\langle g, U|\tilde{T}^-|^{p-1}(I - Q)y\rangle = 0.$$

As y and g are arbitrary, therefore $U|\tilde{T}^-|^{p-1}(I - Q) = 0$. So,

$$\operatorname{Ran}|\tilde{T}^-|^{p-1}(I - Q) \subseteq \operatorname{Ker} U = \operatorname{Ker}|\tilde{T}^-|$$

so that $0 = |\tilde{T}^-|^p(I - Q) = |\tilde{T}^-|(I - Q)$ and hence

$$0 = U|\tilde{T}^-|(I - Q) = \tilde{T}^-(I - Q).$$

Since $g \in \operatorname{Ran} P$ then $g = Px$ for arbitrary x. So from (1) we get

$$0 = \mathscr{R}\langle|\tilde{T}^-|^{p-1}U^*Px, f\rangle.$$

Therefore, $|\tilde{T}^-|^{p-1}U^*P = 0$, hence $|\tilde{T}^-|U^*P = 0$, that is, $(\tilde{T}^-)^*P = 0$. Hence, $P\tilde{T}^- = 0$.

Thus, $\tilde{T}^-(I - Q) = 0 = P\tilde{T}^-$. Therefore, Lemma 1.4 (b) implies that $\tilde{T}^- = T^+$. ■

There is no version of Theorem 3.2 for $0 < p \leq 1$ since the condition that the critical point is invertible would mean that T^+, and hence T, is invertible, so $\|T^-\|_p$ would be constant.

We comment on the hypothesis in Theorems 3.1 (b) and 3.2 that there exists a generalized inverse of T in $\mathscr{C}_p$. In fact, compactness alone of T^- forces that both T and T^* are of finite rank. For if T^-, and hence T $(= TT^-T)$, is compact then, since $\operatorname{Ran} T$ is closed, $\operatorname{Ran} T$ must be finite-dimensional [9, p. 298, Theorem 7.4]; and, as

$$T^* = ((TT^+)T)^* = T^*(TT^+),$$

$\operatorname{Ran} T^*$ is finite-dimensional.

References

1. J.G. Aiken, J.A. Erdos and J.A. Goldstein, *Unitary approximation of positive operators*, Illinois J. Math., vol. 24 (1980), pp. 61–72.
2. R. Bouldin, *Best approximation of a normal operator in the Schatten p-norm*, Proc. Amer. Math. Soc., vol. 80 (1980), pp. 277–282.

3. N. Dunford and J.T. Schwartz, *Linear operators, Part II*, Interscience, New York, 1963.
4. H. Engl and M.Z. Nashed, *New extremal characterisations of generalized inverses of linear operators*, J. Math. Anal. Appl., vol. 82 (1981), pp. 566–586.
5. P.J. Maher, *Partially isometric approximation of positive operators*, Illinois J. Math., vol. 33 (1989), pp. 227–243.
6. R. Penrose, *A generalized inverse for matrices*, Proc. Cambridge Phil. Soc., vol. 51 (1955), pp. 406–413.
7. ______, *On best approximation solutions of linear matrix equations*, Proc. Cambridge Phil. Soc., vol. 52 (1956), pp. 17–19.
8. J.R. Ringrose, *Compact non-self-adjoint operators*, Van Nostrand Rheinhold, London, 1971.
9. A.E. Taylor and D.C. Lay, *Introduction to functional analysis*, 2nd edition, Wiley, New York, 1980.

Middlesex Polytechnic
Trent Park, Bramley Road
London N14 4XS, England