Operators With a Unique Positive Near-Approximant

RICHARD BOULDIN

Communicated by the Editors

§1. Introduction. This paper continues the research began in [4]. Both papers deal with problems suggested by P. R. Halmos in [6]; these problems deal with approximating a given operator with a nonnegative operator. Professor Halmos asked for a characterization of when a given operator has a unique nonnegative operator closest to it.

In this paper we consider two different norms that might be used for measuring the “closeness” of operators. One is the usual operator norm and the other is a topologically equivalent norm which has some computational advantages. In the latter norm we give a complete solution to the above problem.

§2. Preliminaries. By operator we shall mean a bounded linear map of the complex Hilbert space H into itself. If S is a subspace of H then we denote the restriction of T to S by $T|S$; the spectrum of T is $\sigma(T)$; we simply write c when we mean cI with I the identity transformation. To say that T is normaloid means that the numerical radius of T , denoted $w(T)$, is equal to the norm of T , denoted $\|T\|$.

We shall follow Halmos' convention of using “positive operator” as synonymous with “nonnegative operator”; thus a positive operator is not necessarily invertible. We also use Halmo's phrase “positive approximant” to mean the best approximation in operator norm among the positive operators. For the reader's convenience we restate the following results proved by Halmos in [6].

2.1 Theorem. *If $B + iC$ is the usual Cartesian representation for the operator T , then*

$$\inf \{\|T - P\| : P \geq 0\} = \inf \{r : B + (r^2 - C^2)^{1/2} \geq 0\}.$$

The first infimum shall henceforth be denoted $\delta(T)$.

2.2 Theorem. *If $B + iC$ is the usual Cartesian representation for the operator T and if $P_0 = B + (\delta(T)^2 - C^2)^{1/2}$, then P_0 is a positive approximant for T .*

We shall henceforth denote the above positive approximant for any operator T by P_0 and we shall refer to it as the Halmos positive approximant.

We shall use throughout this paper the existence and uniqueness of the Cartesian representation. For any operator T there are two unique self adjoint operators B and C such that $T = B + iC$. The self adjoint operator B can be written as the difference of two positive operators, say $B_+ - B_-$, where B_+ is defined by the equation $2B_+ = B + (B^2)^{1/2}$.

§3. A new point of view. P. R. Halmos proposed that operator theorists solve certain problems of best approximation in the operator norm and, in particular, that we characterize the operators T with the property that T has a unique positive approximant. This problem shall be referred to as the “uniqueness problem”. It now appears that this problem and some related problems are more naturally posed when one uses a norm different from the operator norm but topologically equivalent to the operator norm. If $B + iC$ is the Cartesian representation for the operator T then the norm is simply defined as the square root of the operator norm of $B^2 + C^2$. This norm was implicit in Halmos’ proofs of his basic results in [6] and he explicitly noted the first inequality in (b) of Theorem 3.1. For T we denote this norm by $|||T|||$ and we establish its basic properties in the next theorem.

3.1 Theorem. *The formula $|||T||| = 2^{-1/2} ||T^*T + TT^*||^{1/2}$ defines a norm on $B(H)$ with the following properties:*

- (a) $|||T||| = ||(\operatorname{Re} T)^2 + (\operatorname{Im} T)^2||^{1/2}$,
- (b) $||T|| \geq |||T||| \geq w(T)$ where $w(T)$ is the numerical radius of T ,
- (c) $\inf \{ |||T - P||| : P \geq 0 \} = \delta(T)$,
- (d) $|||T - P_0||| = \delta(T)$.

Proof. The only property of norms which is not obvious is the triangle inequality. To prove this we take a unit vector g and note the following:

$$\begin{aligned}
 |||S||| + |||T||| &\geq 2^{-1/2} (||[S^*S + SS^*]^{1/2}g||^2)^{1/2} + 2^{-1/2} (||[T^*T + TT^*]^{1/2}g||^2)^{1/2} \\
 &= 2^{-1/2} \langle [S^*S + SS^*]g, g \rangle^{1/2} + 2^{-1/2} \langle [T^*T + TT^*]g, g \rangle^{1/2} \\
 &= 2^{-1/2} (||Sg||^2 + ||S^*g||^2)^{1/2} + 2^{-1/2} (||Tg||^2 + ||T^*g||^2)^{1/2}.
 \end{aligned}$$

Using the triangle inequality for the Euclidean 2-space norm, we get

$$\begin{aligned}
 |||S||| + |||T||| &\geq 2^{-1/2} [(||Sg|| + ||Tg||)^2 + (||S^*g|| + ||T^*g||)^2]^{1/2} \\
 &\geq 2^{-1/2} [|(S + T)g|^2 + |(S^* + T^*)g|^2]^{1/2} \\
 &= 2^{-1/2} \langle [(S + T)^*(S + T) + (S + T)(S + T)^*]g, g \rangle^{1/2}.
 \end{aligned}$$

Consequently,

$$\begin{aligned} |||S||| + |||T||| &\geq 2^{-1/2} [w((S+T)^*(S+T) + (S+T)(S+T)^*)]^{1/2} \\ &= 2^{-1/2} |||(S+T)^*(S+T) + (S+T)(S+T)^*|||^{1/2} = |||S+T|||. \end{aligned}$$

The equation of (a) follows from the identity

$$\frac{1}{2}(T^*T + TT^*) = (\operatorname{Re} T)^2 + (\operatorname{Im} T)^2.$$

Let $B = \operatorname{Re} T$ and $C = \operatorname{Im} T$. We shall use (a) to show the second inequality of (b). For any unit vector f we have

$$\begin{aligned} ||(B^2 + C^2)^{1/2}f||^2 &= ||Bf||^2 + ||Cf||^2 \geq (\langle Bf, f \rangle)^2 + (\langle Cf, f \rangle)^2 \\ &= |\langle (B + iC)f, f \rangle|^2 \end{aligned}$$

and this implies that $|||T||| = ||(B^2 + C^2)^{1/2}|| \geq w(T)$. The first inequality of (b) follows directly from the definition of $|||T|||$.

It is apparent from the first inequality of (b) that $\inf \{|||T - P||| : P \geq 0\} \leq \delta(T)$. In order to prove the opposite inequality we recall Theorem 2.1 and we set $r = |||T - P|||$ where $P \geq 0$. Then

$$r^2 = |||T - P|||^2 = ||(B - P)^2 + C^2||$$

which implies $r^2 \geq (B - P)^2 + C^2$

or

$$(r^2 - C^2)^{1/2} \geq |B - P| \geq P - B.$$

Here we use that the square root is a monotone operator function (see Lemma 3.1 [4]). From the above inequality we see that $B + (r^2 - C^2)^{1/2}$ is nonnegative. Now 2.1 implies $r \geq \delta(T)$ and (c) is proved.

Since $||T - P_0|| = \delta(T)$ by 2.2 and $|||T - P_0||| \geq \delta(T)$ by (c), the assertion in (d) follows (b).

From (b) it is immediate that the infimum $\inf \{|||T - P||| : P \geq 0\}$ is assumed by $|||T - P|||$ whenever P is a positive approximant. Motivated by this observation we agree to refer to P as a *positive near-approximant* provided that $|||T - P|||$ assumes the above infimum and we note that any positive approximant is a positive near-approximant. It is a nontrivial fact that not every positive near-approximant is a positive approximant. Now we note a key property of the Halmos positive approximant.

3.2 Corollary. *If P is any positive near-approximant, then $P_0 \geq P$.*

Proof. Because $\delta(T)^2 = |||T - P|||^2 = ||(B - P)^2 + C^2||$ we have

$$\delta^2 - C^2 \geq (B - P)^2$$

or

$$(\delta^2 - C^2)^{1/2} \geq |B - P| \geq P - B$$

or

$$P_0 = B + (\delta^2 - C^2)^{1/2} \geq P \quad \text{as desired.}$$

§4. A solution to the uniqueness problem. The simplicity and completeness of the next two theorems suggest that the uniqueness problem might be profitably reformulated. We characterize the operators T which have a unique positive near-approximant.

4.1 Lemma. *If $R_1, R_2, \dots, R_n$ are positive commuting operators on H , then there is a positive operator R such that*

- (a) $RR_i = R_iR$ for $j = 1, \dots, n$,
- (b) $R \leq R_i$ for $j = 1, \dots, n$,
- (c) $(RH)^- = \bigcap \{(R_iH)^- : j = 1, \dots, n\}$,
- (d) $RH \subset \bigcap \{R_iH : j = 1, \dots, n\}$.

Proof. Let $F(t)$ be the orthogonal projection onto $\text{span } \{E_1(t)H, \dots, E_n(t)H\}^-$ (the bar means topological closure) where $E_i(t)$ is the resolution of identity for R_i , $j = 1, \dots, n$. Clearly $F(a) = 0$ for $a < 0$ and $F(b) = I$ for $b = \max \{\|R_1\|, \dots, \|R_n\|\}$. It is straightforward to see that $s \leq t$ implies $F(s)F(t) = F(s)$.

Take $f \in H$ and let t be a real number; let $\epsilon > 0$. Take $f_1(t)$ from $E_1(t)H$ and $f_j(t)$ from $[E_j(t)H - \text{Span } \{E_1(t)H, \dots, E_{j-1}(t)H\}]$ with $j = 2, \dots, n$ such that

$$\|F(t)f - f_1(t) - \dots - f_n(t)\| < \epsilon/2.$$

Because $f_j(t) \in E_j(t)H$ we have $\lim_{s \rightarrow t+0} E_j(s)f_j(t) = f_j(t)$ and we may take $s_0 \in (t, \infty)$ such that $s \in (t, s_0)$ implies

$$\|f_j(t) - E_j(s)f_j(t)\| < \epsilon/2n.$$

By the definition of orthogonal projection it follows that

$$\begin{aligned} \|F(t)f - F(s)f\| &\leq \|F(t)f - (E_1(s)f_1(t) + \dots + E_n(s)f_n(t))\| \\ &\leq \epsilon/2 + \sum_{j=1}^n \|f_j(t) - E_j(s)f_j(t)\| < \epsilon. \end{aligned}$$

Thus $\lim_{s \rightarrow t+0} F(s)f = F(t)f$. We have now established that $F(t)$ is a resolution of identity and by Theorem 1, p. 36, [1] there is a unique self adjoint operator R associated with $F(t)$. It is routine to see that the spectrum of R is contained in $[0, \infty)$ and consequently it is a positive operator.

It is straightforward to see that

$$F(s)E_j(t) = E_j(t)F(s) \quad \text{for } j = 1, \dots, n.$$

It follows that R commutes with R_j for $j = 1, \dots, n$.

Using integration by parts for the Riemann–Stieltjes integral we get

$$\begin{aligned}
 \langle Rf, f \rangle &= \int_{[0, b]} t \, d\langle F(t)f, f \rangle \\
 &= b \|f\|^2 - \int_{[0, b]} \langle F(t)f, f \rangle \, dt \\
 &\leq b \|f\|^2 - \int_{[0, b]} \langle E_i(t)f, f \rangle \, dt = \int_{[0, b]} t \, d\langle E_i(t)f, f \rangle \\
 &= \langle R_i f, f \rangle \quad \text{for } j = 1, \dots, n.
 \end{aligned}$$

The inequality above follows from

$$\langle F(t)f, f \rangle = \|F(t)f\|^2 \geq \|E_i(t)f\|^2 = \langle E_i(t)f, f \rangle$$

which is established by expanding both sides of the inequality

$$\|f - F(t)f\|^2 \leq \|f - E_i(t)f\|^2.$$

Thus $R \leq R_i$ for $j = 1, \dots, n$.

Since $F(0)H = \text{Span} \{E_1(0)H, \dots, E_n(0)H\}^-$ we see that

$$\begin{aligned}
 (RH)^- &= F((0, \infty))H = [F(0)H]^\perp \\
 &= \bigcap_{i=1}^n [E_i(0)H]^\perp = \bigcap_{i=1}^n (R_i H)^-
 \end{aligned}$$

Because $R \leq R_i$ for $j = 1, \dots, n$ and square root is a monotone operator function (see Lemma 3.1 [4]) we get $R^{1/2} \leq R_i^{1/2}$ for $j = 1, \dots, n$. By the main theorem of [5] it follows that $RH \subset R_i H$ for $j = 1, \dots, n$ and so $RH \subset \bigcap_{i=1}^n R_i H$.

The preceding theorem can be used to show that T fails to have a unique positive approximant if either of the following two subspaces are nontrivial: $(P_0 H)^- \cap (C_+ H)^-$ and $(P_0 H)^- \cap (C_- H)^-$. Here C_+ and C_- are the positive and negative parts, respectively, of the imaginary part of T .

4.2 Theorem. *The normal operator T has a unique positive near-approximant if and only if the closed span of $\ker P_0$ and $\ker (\delta^2 - C^2)$ is all of H .*

Proof. We assume that $\text{Span} \{\ker P_0, \ker (\delta^2 - C^2)\}$ is not dense in H and we show that T has a positive near-approximant other than P_0 . We appeal to Theorem 4.1 to obtain a positive operator A which commutes with $(\delta^2 - C^2)^{1/2}$ and is dominated by $(\delta^2 - C^2)^{1/2}$ and by P_0 . As mentioned in [4] taking squares is a monotone operator function when applied to pairs of commuting positive operators. Thus we have

$$[A - (\delta^2 - C^2)^{1/2}]^2 + C^2 \leq [(\delta^2 - C^2)^{1/2}]^2 + C^2 = \delta^2$$

and so

$$\begin{aligned} |||T - (P_0 - A)|||^2 &= |||A + T - P_0|||^2 \\ &= |||A - (\delta^2 - C^2)^{1/2} + iC|||^2 \leq \delta^2. \end{aligned}$$

Now it suffices to prove that $A \neq 0$ and according to Theorem 4.1 we only need that $\{0\} \neq (P_0 H)^- \cap (\delta^2 - C^2)^{1/2} H^- = (\text{Span } \{\ker P_0, \ker (\delta^2 - C^2)\})^\perp$; this is apparent from our hypothesis.

Let P_1 be a positive near-approximant different from P_0 . Thus

$$\delta^2 = |||T - P_1|||^2 = ||(B - P_1)^2 + C^2||$$

and $\delta^2 \geq (B - P_1)^2 + C^2$. As argued in the third paragraph of the proof of Theorem 3.1 it follows that

$$(\delta^2 - C^2)^{1/2} \geq P_1 - B$$

and

$$(\delta^2 - C^2)^{1/2} \geq B - P_1.$$

Thus $P_0 \geq P_0 - P_1 \geq 0$ and $2(\delta^2 - C^2)^{1/2} \geq P_0 - P_1 \geq 0$ and Closed Span $\{\ker P_0, \ker (\delta^2 - C^2)\}$ is contained in $\ker (P_0 - P_1)$. Since $P_1 \neq P_0$ it must be that Closed Span $\{\ker P_0, \ker (\delta^2 - C^2)\} \neq H$ as desired.

A trivial consequence of Theorem 2.2 is that $\delta \geq \|C\|$ where $B + iC$ is the Cartesian representation of T . In the event that $\delta > \|C\|$ the preceding theorem gives a particularly simple criterion for a unique positive near-approximant—namely $P_0 = 0$.

§5. Similarity to previous results. In [4] we solved the uniqueness problem for several special classes of operators. Here we note that the solution of uniqueness problem for positive near-approximants is exactly the same. A key observation is that the operator norm and the new norm agree for any normaloid operator and consequently any positive near-approximant P is a positive approximant provided $T - P$ is normaloid.

5.1 Lemma. *If $T = B + iC$ (with $B = B^*$, $C = C^*$) and P_0 coincides with the positive part of the real part of T , i.e. $P_0 = B_+$, then P_0 is the only positive near-approximant for T .*

Proof. Since $P_0 = B_+$ and $\delta^2 - C^2 = (B_-)^2$, it is clear that the span of $\{\ker P_0, \ker (\delta^2 - C^2)\}$ is all of H .

5.2 Theorem. *If T is a normal operator, then the following conditions are equivalent:*

- (a) T has a unique positive near-approximant,
- (b) T has a unique positive approximant,
- (c) $P_0 = B_+$
- (d) for some $d \geq 0$ the spectrum of T is contained in $\{z : \operatorname{Re} z \leq 0, |z| = d\} \cup \{z : 0 \leq \operatorname{Re} z \leq \|T\|, \operatorname{Im} z = \pm d\}$.

Proof. It is trivial that (a) implies (b) which implies (c) and the equivalence of (b) and (d) was established in [4]. That (c) implies (a) follows from the lemma and we are done.

5.3 Theorem. *Let $T = B + iC$ (with $B = B^*$, $C = C^*$) be given with $B \geq 0$. The following conditions are equivalent:*

- (a) *T has a unique positive near-approximant,*
- (b) *T has a unique positive approximant,*
- (c) $C^2 = \|C\|^2 I$.

Proof. Clearly (a) implies (b) and the equivalence of (b) and (c) was established by Theorem 4.6 [4]. By Theorem 2.1 $\delta(T) = \|C\|$, and $P_0 = B = B_+$. By Lemma 5.1 we deduce (a).

5.4 Theorem. *The following statements are equivalent:*

- (a) *the zero operator is the unique positive near-approximant of T ,*
- (b) *the zero operator is the unique positive approximant of T ,*
- (c) *T has the form $\delta(T)(-(1 - C^2)^{1/2} + iC)$ where C is a self adjoint contraction.*

Proof. That (a) implies (b) is obvious and that (b) implies (c) follows immediately from Theorem 5.5 in [4]. By that same theorem (c) implies that $P_0 = 0$; thus, Corollary 3.2 implies that the only positive near-approximant is the zero operator.

REFERENCES

1. N. I. AKHIEZER & I. M. GLAZMAN, *Theory of linear operators in Hilbert space, Vol. II* Frederick Unger Publishing Co., New York, 1963.
2. T. ANDO, T. SEKIGUCHI & T. SUZUKI, *Approximation by positive operators*, preprint.
3. R. H. BOULDIN, *The norm continuity properties of square roots*, S.I.A.M. J. Math. Anal. **3** (1972), 206–210.
4. R. H. BOULDIN, *Positive approximants*, Trans. Amer. Math. Soc. **177** (1973), 391–403.
5. R. G. DOUGLAS, *On majorization, factorization and range inclusion of operators in Hilbert space*, Proc. Amer. Math. Soc. **17** (1966), 413–416.
6. P. R. HALMOS, *Positive approximants of operators*, Indiana Univ. Math. J. **21** (1972), 951–961.

University of Georgia

Date communicated: MARCH 5, 1973