Approximation Theory and Matrix Completions

D. Hadwin, D.R. Larson† and D. Timotin

Abstract. A new point of view is introduced in the study of completion problems for partially defined matrices, by relating it to approximation theory in the Banach space of matrices endowed with the operator norm. The approach is interesting especially for completion patterns for which precise results are not possible; several results are obtained in this direction.

Keywords: Partially defined matrices, matrix completion, approximation theory.

2000 Mathematics Subject Classification: 15A60, 65F35, 41A50.

0. Introduction.

The aim of this paper is to relate two distinct areas of analysis which have remained up to now relatively separated. These are, on one side, approximation theory in Banach spaces, and on the other side, completion problems for partially defined matrices. Finding the minimum norm of the completion of a matrix which has only some of the entries specified is an old problem. However, the focus has been mostly in the case when explicit formulae for the minimum completion and of its norm can be obtained. For the (large variety of) patterns for which there does not exist such a nice formula, it is however reasonable, from the point of view of applications, to start by applying an approximation procedure in order to obtain the minimum completion. Such an approximation procedure has best chances to behave well (that is, to converge to the actual solution and to be

† The work of the first and second authors was partially supported by grants from the National Science Foundation.

relatively stable with respect to perturbations of the data) in case the minimum norm completion is unique, and moreover depends continuously on the given data.

It is here that approximation theory in Banach spaces comes into play. Since the completion problem can be reformulated as the distance problem to a given subspace of the space of all matrices on a finite dimensional linear space, it is reasonable to introduce in the study its concepts and results. Although the finite dimensionality assumptions simplify the approximation theory frame up to a certain point, we are still left with some nontrivial questions. Below we try to give answers to some of them; we hope that this is a promising area of research. Some of the facts presented are not new, but, hopefully, a different light is thrown upon them.

1. Approximation theory.

We start by stating the basic notations and results from general approximation theory (see, for instance, [S]).

Let $\mathcal{E}$ be a Banach space, $\mathcal{G} \subset \mathcal{E}$ a linear subspace. If $x \in \mathcal{E}$, then the main problem of approximation theory is the determination of the quantity

$$d_{\mathcal{G}}(x) = \inf\{\|x - y\| \mid y \in \mathcal{G}\}$$

(the distance from x to $\mathcal{G}$) and of the set

$$\mathcal{P}_{\mathcal{G}}(x) = \{y \in \mathcal{G} \mid \|x - y\| = d_{\mathcal{G}}(x)\},$$

(the set of best approximation of x by elements of $\mathcal{G}$).

In this general context, it is possible that $\mathcal{P}_{\mathcal{G}}(x)$ is empty, has one element or several. Since we will work in finite dimensional spaces, the first case cannot occur. The subspaces $\mathcal{G}$ which have the property that $\mathcal{P}_{\mathcal{G}}(x)$ contains exactly one element for any $x \in \mathcal{E}$ are called *Chebyshev* subspaces of $\mathcal{E}$. In case $\mathcal{G}$ is Chebyshev, the uniquely determined element of $\mathcal{P}_{\mathcal{G}}(x)$ is denoted with $p_{\mathcal{G}}(x)$; map $p_{\mathcal{G}} : \mathcal{E} \to \mathcal{G}$ is called the *metric projection* onto $\mathcal{G}$. It is homogeneous, but in general nonlinear.

Even if $\mathcal{G}$ is not Chebyshev, we can define

$$\mathcal{U}(\mathcal{G}) = \{x \in \mathcal{E} \mid \mathcal{P}_{\mathcal{G}}(x) \text{ has one element}\},$$

and consider the metric projection as a map from $\mathcal{U}(\mathcal{G})$ to $\mathcal{G}$. We have the following simple lemma:

Lemma 1.1. *The function $p_{\mathcal{G}}$ is continuous on $\mathcal{U}(\mathcal{G})$.*

Proof. Suppose $x_n, x \in \mathcal{U}(\mathcal{G})$, $x_n \to x$. We have, for any $x \in \mathcal{U}(\mathcal{G})$, $\|p_{\mathcal{G}}(x)\| \leq 2\|x\|$ (since $0 \in \mathcal{G}$); therefore there exists a subsequence (n_k) such that $p_{\mathcal{G}}(x_{n_k})$ converges to some $y \in \mathcal{G}$. Since the distance to $\mathcal{G}$ is continuous on $\mathcal{E}$,

$$\|x - y\| = \lim \|x_{n_k} - p_{\mathcal{G}}(x_{n_k})\| = \lim d_{\mathcal{S}}(x_{n_k}, \mathcal{G}) = d_{\mathcal{S}}(x, \mathcal{G}),$$

whence, x being in $\mathcal{U}(\mathcal{G})$, it follows that $y = p_{\mathcal{G}}(x)$. We have thus shown that any sequence (x_n) contains a subsequence such that $p_{\mathcal{G}}(x_{n_k}) \to p_{\mathcal{G}}(x)$; a standard argument finishes the proof. ∎

Note that $p_{\mathcal{G}}$ is not, in general, uniformly continuous on $\mathcal{U}(\mathcal{G})$; an example will be given in section 3. However, if $\mathcal{G}$ is Chebyshev (and thus $\mathcal{U}(\mathcal{G}) = \mathcal{E}$), the compactness of the unit ball of $\mathcal{E}$ together with the homogeneity of $p_{\mathcal{G}}$ implies its uniform continuity.

2. Spaces of matrices.

We are actually interested in the investigation of approximation problems for spaces of matrices. Denote by $\mathcal{H}_N$ a complex Hilbert space of dimension N, and by $\mathcal{B}(\mathcal{H}_N)$ the algebra of the bounded linear operators on $\mathcal{H}_N$. Of course any orthonormal basis in $\mathcal{H}_N$ produces an identification of $\mathcal{B}(\mathcal{H}_N)$ with $\mathcal{M}_N$, the algebra of square $N \times N$ matrices with complex coefficients. $\mathcal{S}$ will usually be a subspace of $\mathcal{B}(\mathcal{H}_N)$ (equivalently, of $\mathcal{M}_N$).

The approximation problem discussed above becomes then the determination, for $T \in \mathcal{B}(\mathcal{H}_N)$, of the quantity

$$d_{\mathcal{S}}(T) = \inf\{\|T - S\| \mid S \in \mathcal{S}\}.$$

The next result was independently obtained by several authors, including B. Magajna, V. Shulman, and Ileana Ionascu, in response to a question raised in [KL2].

Lemma 2.1. *Suppose $S, T \in \mathcal{B}(\mathcal{H}_N)$, and $\mathcal{S} = \mathbf{C}S$. There exist unit vectors $e, f \in \mathcal{H}_N$, such that $\langle Se, f\rangle = 0$, while $\langle Te, f\rangle = d_{\mathcal{S}}(T)$.*

A related result allows us to obtain easily in the next sections examples of subspaces of $\mathcal{B}(\mathcal{H}_N)$ which are not Chebyshev.

Lemma 2.2. *With the above notations, suppose there are unit vectors $e, f \in \mathcal{H}_N$, such that $\mathcal{S}e \perp f$, and there exists $S \in \mathcal{S}$ with $\|S\| = 1$, such that $Se = S^*f = 0$. Then $\mathcal{S}$ is not Chebyshev.*

Proof. Let $K = f \otimes e$ (that is, $Kh = \langle h, e\rangle f$). Then for any $A \in \mathcal{S}$ we have

$$\|K - A\| \geq |\langle (K - A)e, f\rangle| = \langle f, f\rangle = 1 = \|K\|.$$

Therefore 0 is a best approximant in $\mathcal{S}$. However, $\|K - S\| = 1$, so S is another different approximant. ∎

Chebyshev subspaces are obviously invariant with respect to right or left multiplication with unitaries — that is, $\mathcal{S}$ is Chebyshev iff $U\mathcal{S}$ and/or $\mathcal{S}U$ is Chebyshev.

There is a broader context for this problem in which we are interested. Suppose first that $\mathcal{S} = \mathcal{A}$ is actually a norm closed algebra of operators acting on a non necessarily finite dimensional Hilbert space $\mathcal{H}$. It is customary to denote by $\mathrm{Lat}(\mathcal{A})$ the lattice of closed subspaces of $\mathcal{H}$ which are invariant to all elements of $\mathcal{A}$; conversely, if L is a lattice of subspaces of $\mathcal{H}$, then $\mathrm{Alg}(L)$ is the algebra of all operators which leave the elements of L invariant. $\mathcal{A}$ is then called *reflexive* if $\mathrm{Alg}(\mathrm{Lat}(\mathcal{A})) = \mathcal{A}$.

There is an alternate way of formulating reflexivity. For a closed subspace $M \subset \mathcal{H}$, let us denote by P_M the orthogonal projection onto M. Consider, for any $T \in \mathcal{B}(\mathcal{H})$, the quantity

$$\delta_{\mathcal{A}}(T) = \sup\{\|(I - P_M)TP_M\| \mid M \in \mathrm{Lat}(\mathcal{A})\}.$$

Since $(I - P_M)AP_M = 0$ for any $A \in \mathcal{A}$ and $M \in \mathrm{Lat}(\mathcal{A})$, it follows that $\delta_{\mathcal{A}}(T) \le d_{\mathcal{A}}(T)$. Moreover, $\mathcal{A}$ is reflexive iff $\delta_{\mathcal{A}}(T) = 0$ implies $d_{\mathcal{A}}(T) = 0$. It is called *hyperreflexive* if there exists a positive constant C such that $d_{\mathcal{A}}(T) \le C\delta_{\mathcal{A}}(T)$ for any $T \in \mathcal{B}(\mathcal{H})$.

Hyperreflexivity obviously gives an estimate of the distance to $\mathcal{A}$ in terms of another quantity that might be easier to calculate. A typical (and seminal) example is given by Arveson's result ([A]) that says that $C = 1$ (and thus $d(T, \mathcal{A}) = \delta_{\mathcal{A}}(T)$) for nest algebras. The constant C (called *constant of hyperreflexivity*, or *distance constant*) has been an intense object of study (see, for instance, [D], [DO], [DP], [H], [KL1], [KL2], [L]).

In case we replace the algebra by a subspace $\mathcal{S}$ of $\mathcal{B}(\mathcal{H})$, we have to define ([KL])

$$\delta_{\mathcal{S}}(T) = \sup \|XTY\|,$$

the supremum being taken with respect to all contractions $X, Y \in \mathcal{B}(\mathcal{H})$ such that $XSY = 0$ for any $S \in \mathcal{S}$ (or, equivalently, X, Y are projections with the same property). Again, the subspace is called hyperreflexive if there exists a positive constant C such that $d_{\mathcal{S}}(T) \le C\delta_{\mathcal{S}}(T)$ for any $T \in \mathcal{B}(\mathcal{H})$. Our subspaces will usually be included in $\mathcal{B}(\mathcal{H}_N)$ and thus finite dimensional, so a compactness argument yields immediately hyperreflexivity; one is then interested in the dependence of the distance constant on N.

3. Completion problems.

We come now to the completion problem for partially defined matrices. The latter can be stated as follows: suppose we are only given certain entries of the matrix $A = (a_{ij})_{i,j=1}^n \in \mathcal{M}_N$; that is, we know $a_{ij} = \tilde{a}_{ij}$ only for (i, j) belonging to a certain "pattern" $\Pi \subset$

$\{1, \ldots, n\} \times \{1, \ldots, n\}$. We want to minimize the norm of A subject to these restrictions; that is, to find the quantity

$$\mu = \min\{\|(a_{ij})\| \mid a_{ij} = \tilde{a}_{ij} \text{ for } (i,j) \in \Pi\}$$

(and, of course, also the minimizing matrix). Now, it is obvious that, if we define $\mathcal{S} \subset \mathcal{M}_N$ by

$$\mathcal{S} = \{A = (a_{ij}) \in \mathcal{M}_N \mid a_{ij} = 0 \text{ for } (i,j) \in \Pi\},$$

while $T = (t_{ij})$ is given by

$$t_{ij} = \tilde{a}_{ij} \text{ for} |\ (i,j) \in \Pi \text{ and } t_{ij} = 0 \text{ for } (i,j) \notin \Pi, \tag{1}$$

then

$$\mu = d_{\mathcal{S}}(T).$$

In this case there is a more familiar formula for $\delta_{\mathcal{S}}$. The pattern Π contains certain "fully defined" submatrices; that is, subsets $C, C' \subset \{1, \ldots, N\}$ such that $C \times C' \subset \Pi$. If A is a partially defined matrix corresponding to Π, then $A_{C,C'} = (a_{ij})_{i \in C, j \in C'}$ is completely determined. For T defined by (1), obviously

$$\mu = d_{\mathcal{S}}(T) \geq \max\{\|A_{C,C'}\|\ |\}.$$

Proposition 3.1. *With the above notations,*

$$\delta_{\mathcal{S}}(T) = \max\{\|A_{C,C'}\| \mid C \times C' \subset \Pi\}$$

Proof. For a set $C \in \{1, \ldots, N\}$, denote by P_C the orthogonal projections spanned by the vectors e_i with $i \in C$. If $C \times C' \subset \Pi$, then $P_{C'} S P_C = 0$ for any $S \in \mathcal{S}$; thus

$$\delta_{\mathcal{S}}(T) \geq \max\{\|A_{C,C'}\| \mid C \times C' \subset \Pi\}.$$

To prove the reverse inequality, take two contractions $X, Y \in \mathcal{B}(\mathcal{H})$ with $XSY = 0$ for any $S \in \mathcal{S}$. Let $\xi, \eta \in \mathcal{H}_N$ be unit vectors such that

$$\|XTY\| = \langle XTY\xi, \eta\rangle = \langle TY\xi, X^*\eta\rangle.$$

If $\xi' = \frac{Y\xi}{\|Y\xi\|}$, $\eta' = \frac{X^*\eta}{\|X^*\eta\|}$, and X', Y' are the rank one projections $\xi' \otimes \xi'$ and $\eta' \otimes \eta'$ respectively, then $\|X'SY'\| = 0$ for any $S \in \mathcal{S}$, and $\|X'TY'\| = \|XTY\|$. Take then as S the matricial units $\mathbf{e}_{ij}$, with $(i,j) \in \Pi'$; if $\xi' = (\xi'_1, \ldots, \xi'_N)$ and $\eta' = (\eta'_1, \ldots, \eta'_N)$, we obtain $\xi'_i \eta'_j = 0$.

Define then $C = \{i \mid \xi'_i \neq 0\}$ and $C' = \{i \mid \eta'_i \neq 0\}$. It follows that $(C \times C') \cap \Pi' = \emptyset$, and thus $C \times C' \subset \Pi$. Then obviously

$$\|X'TY'\| \leq \|P_{C'} T P_C\| = \|A_{C,C'}\|,$$

which proves the desired inequality. ∎

However, our interest lies mostly in the cases where there does not exist a "nice" formula for $d_{\mathcal{S}}(T)$ (in particular, $d_{\mathcal{S}}(T) \neq \delta_{\mathcal{S}}(T)$). It is then still interesting, from the point of view of applications, to see when we have uniqueness of the minimizing element. To get a feeling of the possible situations, we will examine a simple, but interesting example.

Example. Consider the completion of a single element for a 2×2 matrix. That is, take $N = 2$ and, identifying $\mathcal{B}(\mathcal{H}_2)$ with $\mathcal{M}_2$, define

$$\mathcal{S} = \left\{ \begin{pmatrix} 0 & 0 \\ 0 & \lambda \end{pmatrix} \mid \lambda \in \mathbf{C} \right\}.$$

A few calculations show that, if $T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $T \in \mathcal{U}(\mathcal{S})$ if and only if $|b|^2 + |c|^2 > 0$ or $a = b = c = 0$. One can write the precise formulae

$$d(T, \mathcal{S}) = \max\{|a|^2 + |b|^2, |a|^2 + |c|^2\}$$

and

$$p(T) = \begin{cases} \begin{pmatrix} 0 & 0 \\ 0 & d - \frac{c\bar{a}}{\bar{b}} \end{pmatrix} & \text{if } |c| \leq |b| \text{ and } b \neq 0, \\ \begin{pmatrix} 0 & 0 \\ 0 & d - \frac{b\bar{a}}{\bar{c}} \end{pmatrix} & \text{if } |b| \leq |c| \text{ and } c \neq 0, \\ T & \text{if } a = b = c = 0. \end{cases}$$

Therefore, if, say, $T_n = \begin{pmatrix} 1 & 1/n \\ t/n & 0 \end{pmatrix}$ for $0 < t \leq 1$, then $p(T_n) = \begin{pmatrix} 0 & 0 \\ 0 & -t \end{pmatrix}$. If $n \to \infty$, then $T_n \to \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, while $p(T_n)$ is constant in n, but depends on t. Thus, p behaves badly when T approaches $\mathcal{B}(\mathcal{H}_N) \setminus \mathcal{U}(\mathcal{S})$; in particular, it is not uniformly continous on $\mathcal{U}(\mathcal{S})$.

Let us remark also that in this case $\mathcal{U}(\mathcal{S})$ is dense in $\mathcal{B}(\mathcal{H}_N)$. Some other instances of this phenomenon will appear in the next sections.

4. One-dimensional subspaces.

The first case that will be treated is that of one-dimensional subspaces of $\mathcal{B}(\mathcal{H}_N)$ (that is, of spaces of the type $\mathbf{C}K$, with $0 \neq K \in \mathcal{B}(\mathcal{H}_N)$). The result obtained also covers the case of their annihilators.

Theorem 4.1. *Suppose $0 \neq K \in \mathcal{B}(\mathcal{H}_N)$. The following are equivalent:*

(i) *K is invertible;*

(ii) *$\mathbf{C}K$ is Chebyshev;*

(iii) $K_\perp = \{S \in \mathcal{B}(\mathcal{H}_N) \mid \operatorname{tr}(SK) = 0\}$ *is Chebyshev.*

Proof. Obviously we may suppose that $\operatorname{tr}(K) = 1$ and (by using the invariance to unitaries) that $K \geq 0$. Suppose first that K is not invertible. We may then find $e, f \in \mathcal{H}_N$ with $Ke = K^*f = 0$. Applying Lemma 2.2, it follows that $\mathbf{C}K$ is not Chebyshev; so (ii)$\Rightarrow$(i) is true. Also, if we define $\phi : \mathcal{B}(\mathcal{H}_N) \to \mathbf{C}$ by $\phi(T) = \operatorname{tr}(TK)$, then $\|\phi\| = 1$ and $K_\perp = \ker\phi$. Hence, for every $T \in \mathcal{B}(\mathcal{H}_N)$, $\operatorname{dist}(T, K_\perp) = |\phi(T)|$. Thus $T - \phi(T)I$ is a best approximant of T in $K_\perp$. If $T \notin K_\perp$ (so $\phi(T) \neq 0$) and e is a unit vector in $\ker K$, then $T - \phi(T)I + \phi(T)(e \otimes e)$ is a different best approximant. Thus (iii)$\Rightarrow$(i) is proved.

Suppose now that K is invertible. Fix $T \in \mathcal{B}(\mathcal{H}_N)$, and let λK be a best approximant of T in $\mathbf{C}K$. If e, f are the unit vectors given by Lemma 2.1, then

$$\|T - \lambda K\| = \langle Te, f\rangle = \langle (T - \lambda K)e, f\rangle,$$

whence $(T - \lambda K)e = f$. Then $\langle Ke, f\rangle = 0$ implies $\langle (T - \lambda K)e, Ke\rangle = 0$. Since K is invertible, $Ke \neq 0$, and therefore $\lambda = \frac{\langle Te, Ke\rangle}{\langle Ke, Ke\rangle}$. Thus λ is uniquely determined, and consequently $\mathbf{C}K$ is Chebyshev.

Turning now to $K_\perp$, suppose $T - \phi(T)I + B$ is a best approximant for T in $K_\perp$. Then $\operatorname{tr}(BK) = 0$. By choosing a basis with respect to which K is diagonal (remember $K \geq 0$ and $\operatorname{tr} K = 1$), it follows that a convex combination of the diagonal entries of B is 0. On the other hand,

$$|\phi(T)| = \operatorname{dist}(T, K_\perp) = \|T - (T - \phi(T)I + B)\| = \|\phi(T)I - B\|.$$

Hence each of the diagonal entries of B must be in the disk centered at $\phi(T)$ with radius $|\phi(T)|$. Since this disk contains 0 only as an extreme point, all of the diagonal entries of B must be 0. But then $\|\phi(T)I - B\| = |\phi(T)|$ implies that all the nondiagonal entries of B must be 0. Therefore $B = 0$, and it follows that $K_\perp$ is Chebyshev. ∎

The equivalence (i)$\Leftrightarrow$(ii) appears already in [R]. Note that the space $K_\perp$ is also considered with the operator norm, not with the dual trace class norm.

5. The case $N = 2$.

It is possible to determine all the Chebyshev subspaces of $\mathcal{B}(\mathcal{H}_2)$. Theorem 4.1 has already settled the cases $\dim \mathcal{S} = 1$ and $\dim \mathcal{S} = 3$, so we are left only with $\dim \mathcal{S} = 2$.

Lemma 5.1. *If $\mathcal{S}$ is a 2-dimensional subspace of $\mathcal{B}(\mathcal{H}_2)$ that contains an invertible operator, then there are either one or two linearly independent operators of rank 1 in $\mathcal{S}$.*

Proof. Suppose $A \in \mathcal{S}$ is invertible. If $B \in \mathcal{S}$ is not a multiple of A, and $\lambda \in \sigma(A^{-1}B)$, then

$$K = \lambda A - B = A(\lambda - A^{-1}B)$$

is noninvertible and non zero, and therefore has rank 1. We may find (possibly different) bases in the domain and the range of the operators, such that

$$K = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

with $ad - bc \neq 0$. Then

$$\det(\mu K + A) = \begin{vmatrix} \mu + a & b \\ c & d \end{vmatrix} = \mu d + ad - bc;$$

therefore $\det(\mu K + A) = 0$ has one solution if $d \neq 0$, and no solution if $d = 0$. The lemma is proved. ∎

Remark. In the proof it has been shown that, if there exists (up to a scalar multiple) a single operator $K = x \otimes y$ of rank 1 in $\mathcal{S}$ (the notation means that $K\xi = \langle \xi, x \rangle y$), and $e \perp x$, $f \perp y$, then $\langle Be, f \rangle = 0$ for any $B \in \mathcal{S}$.

Theorem 5.2. *If $\mathcal{S}$ is a 2-dimensional subspace of $\mathcal{M}_2$, then $\mathcal{S}$ is Chebyshev iff $\mathcal{S}$ contains two linearly independent operators of rank 1.*

Proof. By Lemma 5.1, the statement is equivalent to the assertion that $\mathcal{S}$ is not Chebyshev iff either it consists only of non-invertible operators, or it contains only one (up to scalar multiples) operator of rank 1.

If $\mathcal{S}$ does not contain invertible operators, one can find a basis such that either $\mathcal{S}$ or $\mathcal{S}^*$ is $\left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \mid a, b \in \mathbf{C} \right\}$. Consider, for instance, the first case. If $T = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, then it has as best approximants in $\mathcal{S}$ all matrices $\begin{pmatrix} 0 & \lambda \\ 0 & 0 \end{pmatrix}$ with $|\lambda| \leq 1$, and therefore $\mathcal{S}$ is not Chebyshev.

Suppose now that $\mathcal{S}$ contains (up to a scalar multiple) a single operator of rank 1 $K = x \otimes y$. If, as above, $e \perp x$ and $f \perp y$, with $\|e\| = \|f\| = 1$, and $T = e \otimes f$, then $\|T\| = 1$, while the remark after the proof of Lemma 5.1 shows that for any $B \in \mathcal{S}$

$$1 = \langle Te, f \rangle = \langle (T - B)e, f \rangle \leq \|T - B\|.$$

Therefore, $d(T, \mathcal{S}) = 1$. However, $\|T - \mu K\| = 1$ for sufficiently small μ; therefore T has not a unique best approximant, and $\mathcal{S}$ is not Chebyshev.

If $\mathcal{S}$ is not Chebyshev, let $T \in \mathcal{M}_2$ be an operator that has not a unique best approximant in $\mathcal{S}$. If $\|T - K_1\| = \|T - K_2\| = d_{\mathcal{S}}(T)$, $K_1 \neq K_2$, then any convex combination of

K_1 and K_2 is a best approximant for T. If $T' = T - \frac{K_1+K_2}{2}$ and $K = \frac{K_2-K_1}{2}$, then μK is a best approximant for T' for all $-1 \le \mu \le 1$. We have then $d_{\mathcal{S}}(T') = d_{\mathbf{C}K}(T') = \|T'\|$; let us also assume, for simplicity, that this last quantity is equal to 1.

Suppose $\mathcal{S}$ contains an invertible operator A. We have

$$\|T'\| \ge d_{\mathbf{C}A}(T') \ge d_{\mathcal{S}}(T') = 1 = \|T'\|$$

All terms are consequently equal, and therefore $d_{\mathbf{C}A}(T') = 1$. By Lemma 2.1 there exist unit vectors e, f, such that $\langle Ae, f\rangle = 0$ and $\langle Te, f\rangle = d(T', \mathbf{C}A) = \|T'\| = 1$.

Complete then e and f to bases $\{e, e'\}$ and $\{f, f'\}$ in the domain and the range of the operators respectively; we will write the matrices of the different operators with respect to these bases. Then $A = \begin{pmatrix} 0 & b \\ c & d \end{pmatrix}$, with $bc \neq 0$, and $T' = \begin{pmatrix} 1 & 0 \\ 0 & t \end{pmatrix}$. Moreover, since $\|T' - \mu K\| = 1$ for all $-1 \le \mu \le 1$, it follows that $K = \begin{pmatrix} 0 & 0 \\ 0 & k \end{pmatrix}$.

Since $\mathcal{S}$ is spanned by A and K, it is easy to see that the only operators of rank 1 in $\mathcal{S}$ are multiples of K. Thus, if $\mathcal{S}$ is not Chebyshev, either it contains only rank 1 operators, or it contains a single (up to scalar multiple) operator of rank 1. The theorem is proved. ∎

6. The diagonal algebra; the case $N = 3$.

As noted in the introduction, the distance problem has been intensively studied in the case when $\mathcal{S}$ is actually a subalgebra of $\mathcal{B}(\mathcal{H}_N)$. One of the important examples is the subalgebra $\mathcal{D}_N$ of operators that are diagonal with respect to a certain basis, which will be fixed during this section. Except in the trivial cases $N = 1$ and $N = 2$), it can be shown that the distance constant C_N satisfies $1 < C_N \le 2$ (see, for instance, [D]). In the case $N = 2$, we may apply Theorem 5.2 to conclude that $\mathcal{D}_2$ is a Chebyshev subspace of $\mathcal{B}(\mathcal{H}_N)$.

We will investigate now the case $N = 3$. We will need the following fact, sometimes known as Parrott's lemma (see [P]); it expresses the fact that, if p, q are positive integers, then $\mathcal{M}_p$ embedded in $\mathcal{M}_{p+q}$ in the upper left corner has distance constant 1.

Lemma 6.1. *If A, B, C are matrices of corresponding dimensions $p \times q$, $q \times p$ and $q \times q$, then*

$$\min_X \left\| \begin{pmatrix} X & A \\ B & C \end{pmatrix} \right\| = \max\{ \left\| \begin{pmatrix} A \\ C \end{pmatrix} \right\|, \|(B \quad C)\| \}.$$

Since we will have to do certain detailed computations, we will use ad-hoc notations for 3×3 matrices. A typical $T \in \mathcal{B}(\mathcal{H}_3)$ will be denoted by

$$T = \begin{pmatrix} \xi & a & b \\ c & \eta & d \\ e & f & \zeta \end{pmatrix}.$$

Thus, in this case

$$d(T, \mathcal{D}) = \min\Big\{ \left\| \begin{pmatrix} x & a & b \\ c & y & d \\ e & f & z \end{pmatrix} \right\| \mid x, y, z \in \mathbf{C} \Big\},$$

while, according to Proposition 3.1,

$$\delta_{\mathcal{D}}(T)^2 = \max\{|a|^2 + |b|^2, |c|^2 + |d|^2, |e|^2 + |f|^2, |c|^2 + |e|^2, |a|^2 + |f|^2, |b|^2 + |d|^2\}.$$

It is simple to give an example of a T for which $d_{\mathcal{D}}(T) > \delta_{\mathcal{D}}(T)$ (take all entries equal to 1). Thus, the algebra $\mathcal{D}$ has indeed constant of hyperreflexivity strictly greater than 1; it is actually equal to $\sqrt{3/2}$ ([DO]). Neither is it a Chebyshev subspace of $\mathcal{B}(\mathcal{H}_3)$ (see Section 8). There is, however, still an interesting phenomenon connected to this fact; it is given by the following result.

Theorem 6.2. *If $d_{\mathcal{D}}(T) > \delta_{\mathcal{D}}(T)$, then T has a unique approximant in $\mathcal{D}$.*

Proof. From Lemma 6.1 it follows that

$$d_{\mathcal{D}}(T) = \min_{x,y} \max\{\|A(x,y)\|, \|A'(x,y)\|\},$$

where

$$A = \begin{pmatrix} x & a & b \\ c & y & d \end{pmatrix}, \quad A' = \begin{pmatrix} x & a \\ c & y \\ e & f \end{pmatrix}.$$

Now, $\|A\| \le r$ is equivalent to $r^2 I_2 - AA^* \ge 0$, or

$$\begin{pmatrix} r^2 - |x|^2 - |a|^2 - |b|^2 & -x\overline{c} - a\overline{y} - b\overline{d} \\ -\overline{x}c - \overline{a}y - \overline{b}d & r^2 - |c|^2 - |y|^2 - |d|^2 \end{pmatrix} \ge 0.$$

By writing that the trace and the determinant of this last matrix are positive, we obtain the two inequalities:

$$\tau_r(x, y) := 2r^2 - |x|^2 - |a|^2 - |b|^2 - |c|^2 - |y|^2 - |d|^2 \ge 0, \tag{2}$$

$$\Delta_r(x, y) := (r^2 - |x|^2 - |a|^2 - |b|^2)(r^2 - |c|^2 - |y|^2 - |d|^2) - |x\overline{c} + a\overline{y} + b\overline{d}|^2 \ge 0. \tag{3}$$

Note that (2) and (3) imply the two inequalities:

$$r^2 - |x|^2 - |a|^2 - |b|^2 \geq 0, \quad r^2 - |c|^2 - |y|^2 - |d|^2 \geq 0 \tag{4}$$

Similarly, the inequality $\|A'\| \leq r$ is equivalent to the two inequalities

$$\tau_r'(x,y) = 2r^2 - |x|^2 - |c|^2 - |e|^2 - |a|^2 - |y|^2 - |f|^2 \geq 0 \tag{5}$$

$$\Delta_r'(x,y) = (r^2 - |x|^2 - |c|^2 - |e|^2)(r^2 - |a|^2 - |y|^2 - |f|^2) - |x\overline{a} + c\overline{y} + e\overline{f}|^2 \geq 0 \tag{6}$$

Denote by $G(r) = \{(x,y) \in \mathbf{C}^2 \mid (2) \text{ and } (3) \text{ are satisfied}\}$ and $G'(r) = \{(x,y) \in \mathbf{C}^2 \mid (4) \text{ and } (5) \text{ are satisfied}\}$. We have $(x,y) \in G(r)$ iff $\|A\| \leq r$, and $(x,y) \in G'(r)$ iff $\|A'\| \leq r$; thus $G(r)$ and $G'(r)$ are (possibly empty) convex sets for any $r \geq \delta_{\mathcal{D}}(T)$, and

$$d(T,\mathcal{D}) = \min\{r \geq \delta_{\mathcal{D}}(T) \mid G(r) \cap G'(r) \neq \emptyset\}$$

In all subsequent statements we assume that $d(T,\mathcal{D}) > \delta_{\mathcal{D}}(T)$. We will now state and prove some intermediate results.

Lemma 6.3. *Suppose that (x_1,y_1), (x_2,y_2) are two distinct pairs of complex numbers such that the segment $I = [(x_1,y_1),(x_2,y_2)] \subset \partial G(r)$. Then either $x_1 = x_2$ and $a = 0$, or $y_1 = y_2$ and $c = 0$.*

Proof. If $(x,y) \in \partial G(r)$, then at least one of the inequalities (2) and (3) have to be an equality; thus, either $\tau_r(x,y) = 0$ or $\Delta_r(x,y)=0$. If $\tau_r(x,y) = 0$, the inequalities (4) become equalities. Therefore the first term of $\Delta_r(x,y)$ is 0. Then $\Delta_r(x,y) \geq 0$ implies actually $\Delta_r(x,y) = 0$, and we may suppose that this last equality is fulfilled for all elements of I.

Now, if we denote $z_t = (tx_1 + (1-t)x_2, ty_1 + (1-t)y_2)$, then $I \subset \partial G(r)$ implies $\Delta_r(z_t) = 0$ for $0 \leq t \leq 1$. Since the function $t \mapsto \Delta_r(z_t)$ is a polynomial in t, it should vanish identically. This polynomial has its leading coefficient (of degree 4) equal to $|x_1 - x_2|^2|y_1 - y_2|^2$. One of the factors in this last product has therefore to vanish.

Suppose then $x_1 = x_2$ (the remaining case is treated analogously). We have then, accordingly to the definition of Δ_r

$$(r^2 - |x_1|^2 - |a|^2 - |b|^2)(r^2 - |c|^2 - |ty_1 + (1-t)y_2|^2 - |d|^2) \\ - |x_1\overline{c} + a\overline{(ty_1 + (1-t)y_2)} + b\overline{d}|^2 = 0 \tag{7}$$

The left hand side is a polynomial of degree 2 in t, which should vanish identically. In particular, this should be true about its leading coefficient, which is $-((r^2 - |x_1|^2 - |a|^2 - |b|^2) + |a|^2)|y_1 - y_2|^2$. By hypothesis, $y_1 \neq y_2$; thus,

$$r^2 - |x_1|^2 - |a|^2 - |b|^2 + |a|^2 = 0.$$

Since, by (4), $r^2 - |x_1|^2 - |a|^2 - |b|^2 \geq 0$, it follows that $a = 0$. ∎

The computations at the end of the proof have in fact shown that in the considered case we have

$$|x_1|^2 = r^2 - |b|^2 \tag{8}$$

and, by (7),

$$x_1\bar{c} + b\bar{d} = 0 \tag{9}$$

Note that it may indeed happen that $\partial G(r)$ contains an interval. As an example, take $a = 0$, $b = 2$, $c = 1$, $d = -1$, $r = 2\sqrt{2}$. Then $\partial G(r)$ contains all pairs (x, y) with $x = 2$, $|y|^2 \leq 6$.

Lemma 6.4. *If $\partial G(r)$ contains a nontrivial segment $I = [(x_1, y_1), (x_2, y_2)]$, then the interior of $G(r)$ is nonempty.*

Proof. According to lemma 6.3, we may assume that $x_1 = x_2$ and $a = 0$ and (8) and (9) are satisfied. If $bd = 0$, the second term in equation (3) vanishes. Since we have $r > \delta_{\mathcal{D}}(T)$, it is immediate that a neighborhood of the origin in $\mathbf{C}^2$ is included in $G(r)$.

Suppose $bd \neq 0$; note that this implies also $c \neq 0$, and $x_1 = -\frac{b\bar{d}}{\bar{c}}$. If the argument of $\frac{b\bar{d}}{\bar{c}}$ is s, write $x = we^{is}$, $w \in \mathbf{R}$. Define then the function

$$\phi(w) = (r^2 - w^2 - |b|^2)(r^2 - |c|^2 - |d|^2) - (|c|w + |bd|)^2$$

This is a function of the real argument w, which vanishes for $w_0 = -\frac{|bd|}{|c|} < 0$. Also, we have $\phi'(w_0) = 2\frac{|bd|}{|c|}(r^2 - |c|^2 - |d|^2) > 0$ (again since $r > \delta_{\mathcal{D}}(T)$). Thus, there exists w_1, $w_0 < w_1 < 0$, such that $\phi(w_1) > 0$. But then we have $\tau_r(w_1 e^{is}, 0) > 0$ and $\Delta_r(w_1 e^{is}, 0) > 0$; therefore a whole neighborhood of the pair $(w_1 e^{is}, 0)$ is contained in $G(r)$. ∎

Obviously lemmas 6.2 and 6.3 are valid for $G'(r)$ as well (with the obvious modifications of notation).

Lemma 6.5. *$H(r) = G(r) \cap G'(r)$ is either empty, or one point, or has interior points.*

Proof. Suppose that $H(r)$ has more than one point, but no interior points. Therefore either $H(r) \subset \partial G(r)$ or $H(r) \subset \partial G'(r)$. Actually, it has to belong to both of them; indeed, if, say, $H(r) \subset \partial G(r)$, then, by lemma 2, $G(r)$ should have a nonempty interior. If $H(r) \not\subset \partial G'(r)$, that is, $H(r)$ is included in the interior of $G'(r)$, then $\partial G(r)$ intersects the interior of $G'(r)$; and therefore the same is true about the interior of $G(r)$, contradicting the assumption that $H(r)$ has no interior.

Now, suppose an interval $I = [(x_1, y_1), (x_2, y_2)] \subset H(r)$. Since $I \subset \partial G(r)$, we have to be in one of the possibilities allowed by lemma 1. Thus, assume that $x_1 = x_2$ and $a = 0$. Also, $I \subset \partial G'(r)$; since we already know that $x_1 = x_2$, the corresponding alternative for $G'(r)$ implies that $c = 0$. But then the choice $x = y = 0$ implies that $d_{\mathcal{D}}(T) = \delta_{\mathcal{D}}(T) = \max\{|e|^2 + |f|^2, |b|^2 + |d|^2\}$, contrary to the original assumption $d_{\mathcal{D}}(T) > \delta_{\mathcal{D}}(T)$. ∎

We may finish now the proof of the theorem. If $r < d_{\mathcal{D}}(T)$, then $G(r) \cap G'(r) = \emptyset$. It follows then that $H(d_{\mathcal{D}}(T))$ cannot have interior points: since $\tau_r(x, y)$ and $\Delta_r(x, y)$ depend continuously of r, we would obtain an open set contained in their intersection for $r < d_{\mathcal{D}}(T)$. On the other hand, the definition of $H(d_{\mathcal{D}}(T))$ implies that it is not empty; therefore, by lemma 6.4, it has to contain a single pair(x_0, y_0). Thus, there exist uniquely determined x_0, y_0 for which there exists z_0 such that $\|M(x_0, y_0, z_0)\| = d_{\mathcal{D}}(T)$. But the original problem is symmetric in x, y, z; thus, z_0 is also unique. The proposition is proved. ∎

7. Some general results.

Proposition 7.1. *If $\mathcal{S}$ is a nontransitive linear subspace of $\mathcal{B}(\mathcal{H}_N)$ and $\dim \mathcal{S} \geq 2N$, then $\mathcal{S}$ is not Chebyshev.*

Proof. Since $\mathcal{S}$ is not transitive, one can take a unit vector $f \perp \mathcal{S}(\mathcal{H}_N)$. Fix any unit vector $e \in \mathcal{H}_N$; obviously $\langle \mathcal{S}e, f \rangle = 0$. Also, the map $S \mapsto (Se, S^* f)$ maps $\mathcal{S}$ into a vector space of dimension $2N - 1$, and thus has a nontrivial kernel. Lemma 2.2 finishes then the proof. ∎

Proposition 7.2. *Suppose that $\mathcal{A} \subset \mathcal{B}(\mathcal{H}_N)$ is a unital subalgebra, which is Chebyshev as a subspace. Then either $\mathcal{A} = \{0\}$, $\mathcal{A} = \mathbf{C}I$, $\mathcal{A} = \mathcal{B}(\mathcal{H}_N)$, or $\mathcal{A} = \operatorname{sp}\{P, I\}$ for some idempotent $P \in \mathcal{B}(\mathcal{H}_N)$.*

Proof. Assume $\mathcal{A}$ is none of $\{0\}, \mathbf{C}I$ or $\mathcal{B}(\mathcal{H}_N)$. Since $\mathcal{A} \neq \mathcal{B}(\mathcal{H}_N)$, it follows from Burnside's theorem that $\mathcal{A}$ has a nontrivial invariant subspace M. Then relative to the decomposition $\mathcal{H}_N = M \oplus M^\perp$ every $A \in \mathcal{A}$ has an operator matrix of the form

$$A = \begin{pmatrix} \alpha(A) & \beta(A) \\ 0 & \gamma(A) \end{pmatrix}.$$

We will first show that $\mathcal{A}$ contains no nonzero nilpotents. Suppose $A \in \mathcal{A}$, $A \neq 0$, and A is nilpotent. Then $\alpha(A)$ and $\gamma(A)$ are both nilpotent. Hence we can choose a unit vector $e \in \ker \alpha(A)$ and a unit vector $f \in \ker(\gamma(A)^*)$. Since $e \in M$ and $f \in M^\perp$, we have $\langle Te, f \rangle$ for any $T \in \mathcal{A}$. Also, $Ae = A^* f = 0$; by Lemma 2.2 we see that $\mathcal{A}$ cannot be Chebyshev. Hence A contains no nonzero nilpotents.

It follows then from Wedderburn's theorem that $\mathcal{A}$ is semisimple and commutative, and hence simultaneously diagonalizable. Thus, there exists a basis $\{e_1, \ldots, e_N\}$ of $\mathcal{H}_N$ with respect to which $\mathcal{A}$ is upper triangular.

Suppose now that the map $\psi(A) = (\langle Ae_1, e_1 \rangle, \langle Ae_N, e_N \rangle)$ is not injective on $\mathcal{A}$. If $A \in \ker \psi$, then for every $T \in \mathcal{A}$, $\langle Te_1, e_N \rangle = 0$, and $Ae_1 = A^* e_N = 0$, which, by Lemma

2.2, violates the fact that $\mathcal{A}$ is Chebyshev. Hence ψ is injective, so $\dim \mathcal{A} \leq 2$. Since $I \in \mathcal{A}$, and $\mathcal{A} \neq \mathbf{C}I$, we must have $\dim \mathcal{A} = 2$. But $\mathcal{A}$ contains no nonzero nilpotents, and therefore there exists an idempotent P, $0 \neq P \neq I$, such that $\mathcal{A} = \mathrm{sp}\{I, P\}$. ∎

8. Well-posedness of the problem.

As noted above, if the subspace $\mathcal{S}$ of $\mathcal{B}(\mathcal{H}_N)$ is Chebyshev, then the metric projection $p_{\mathcal{S}}$ is continuous on $\mathcal{S}$. Even if $p_{\mathcal{S}}$ and $d_{\mathcal{S}}$ are not given by explicit "nice" formulae, this case is rather convenient from the point of view of applications. Basically, the uniqueness of the solution suggests that a numerical procedure that provides an approximate solution has good chances of being convergent. We may then give the following definition: the optimization problem is *well-posed* in $T \in \mathcal{B}(\mathcal{H}_N)$ iff T is in the interior of $\mathcal{U}(\mathcal{S})$. Now, an immediate consequence of Lemma 1.1 is that if the optimization problem is well-posed in T, then $p_{\mathcal{S}}$ is uniformly continuous on a neighborhood of T. As follows from the discussion in the introduction, the condition of well-posedness could be considered as a convenient assumption for the stability of a numerical procedure that computes the solution.

If $\mathcal{S}$ is Chebyshev, then the optimization problem is well-posed in any point of $\mathcal{B}(\mathcal{H}_N)$. However, one of the consequences of the above results, especially of Section 7, show that "most" of the subspaces are not Chebyshev. As is shown by the example in Section 3, we cannot escape pathological situations. The most we can hope is that these are somehow "exceptional"; a way of making this assertion precise is the following definition. We will call a subspace $\mathcal{S} \subset \mathcal{B}(\mathcal{H}_N)$ *almost Chebyshev* if $\mathcal{U}(\mathcal{S})$ contains an open dense subset of $\mathcal{B}(\mathcal{H}_N)$. Let us examine some particular cases, studied in the above sections.

8.1. In the case of one-dimensional subspaces of $\mathcal{B}(\mathcal{H}_N)$, we have to consider $\mathcal{S} = \mathbf{C}K$ with K non-invertible. If $T \in \mathcal{B}(\mathcal{H}_N)$ does not have a unique approximant in $\mathcal{S}$, it follows that $\|T - tK\| = d_{\mathcal{S}}(T) = d$ is constant for t belonging to a whole interval. It follows that

$$\det((d^2 I - (T - tK)(T^* - tK^*)) = 0,$$

which is a polynomial in t, has to be identically 0. In particular, this has to happen for its free term, which is $\det(d^2 I - TT^*)$, and with its highest degree term, which is $\det(d^2 I - KK^*)$. Thus, d is a common singular value for T and K; the set of T for which this does not happen is open and dense.

8.2. Consider now $N = 2$ and $\dim \mathcal{S} = 2$. If $\mathcal{S} = \left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \mid a, b \in \mathbf{C} \right\}$, then $\mathcal{U}(\mathcal{S}) = \mathcal{S}$; thus $\mathcal{S}$ is not almost Chebyshev. Indeed, suppose $T = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$. It is immediate that $d_{\mathcal{S}}(T) = |z|^2 + |t|^2$. Since $\left\| \begin{pmatrix} \bar\lambda t & -\lambda \bar z \\ z & t \end{pmatrix} \right\| = |z|^2 + |t|^2$ for $0 \leq \lambda \leq 1$, we have $d_{\mathcal{S}}(T) = \|T - T_\lambda\|$, where $T' = \begin{pmatrix} x - \lambda \bar t & y + \lambda \bar z \\ 0 & 0 \end{pmatrix}$. Therefore, if $|z|^2 + |t|^2 > 0$, we have different best approximants for each value of λ.

The remaining case to be studied is the one when $\mathcal{S}$ contains a single (up to a scalar multiple) operator of rank 1. The remark after Lemma 5.1 shows that in this case we may find bases such that

$$\mathcal{S} = \left\{ \begin{pmatrix} a & b \\ kb & 0 \end{pmatrix} \mid a, b \in \mathbf{C} \right\}$$

for some fixed value of $k \neq 0$. If $T = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$, then the same argument used for the example in Section 3 shows that

$$d_{\mathcal{S}}(T) = \inf_{b \in \mathbf{C}} \max\{|t|^2 + |y - b|^2, |t|^2 + |z - kb|^2\}.$$

If $z \neq ky$, then the value of b which realizes the minimum is uniquely determined, and so is then a for which $d_{\mathcal{S}}(T) = \|T - S\|$. Thus, the space is almost Chebyshev.

8.3. Consider now $\mathcal{S} = \mathcal{D}_3$. We have to look closely in the proof of Theorem 6.2. Lemmas 6.5 and 6.3 show that, in order to have nonuniqueness of the best approximation in $\mathcal{D}_3$ for an operator T, it is necessary that at least one of the entries of T be null. It follows then that $\mathcal{D}_3$ is indeed almost Chebyshev.

As noted above, almost Chebyshev subspaces are interesting from the point of view of numeric procedures for finding the best approximation. If the approximation problem is well-posed in T, it is reasonable to expect a good behaviour of the approximation scheme. For almost Chebyshev subspaces $\mathcal{S}$ this is the case for "most" T's. We believe the investigation of almost Chebyshev subspaces and of points of well-posedness for the approximation problem to deserve a more detailed study.

REFERENCES

[A] W. Arveson: Interpolation problems in nest algebras, *J. Functional Analysis* **20** (1975), 208–233.

[D] K.R. Davidson: Finite dimension problems in operator theory, *The Gohberg Anniversary Collection*, vol.1, Birkhäuser Verlag, Basel-Boston-Berlin, 187–201.

[DO] K.R. Davidson and M.S. Ordower: Some exact distance constants, *Linear Algebra Appl.* **208/209** (1994), 37–55.

[DP] K.R. Davidson and S.C. Power: Failure of the distance formula, *J. London Math. Soc.*(2) **32** (1985), 157–165.

[H] D. Hadwin: A general view of reflexivity, *Trans. Amer. Math. Soc.* **344** (1994), 325–360.

[KL1] J. Kraus and D.R. Larson: Some applications of a technique for constructing reflexive operator algebras, *J. Operator Theory* **13** (1985), 227–236.

[KL2] J. Kraus and D.R. Larson: Reflexivity and distance formulae, *Proc. London Math. Soc.* (3) **53** (1986), 340–356.

[L] D.R. Larson: Annihilators of operator algebras, in *Invariant subspaces and other topics (Timişoara/Herculane, 1981)*, pp. 119–130, Operator Theory: Adv. Appl., 6, Birkhäuser, Basel-Boston, Mass., 1982.

[M] B. Magajna: On the distance to finite-dimensional subspaces in operator algebras, *J. London Math. Soc.* (2) **47** (1993), 516–532.

[P] S. Parrott: On a quotient norm and the Sz.-Nagy–Foias lifting theorem, *J. Functional Analysis* **30** (1978), 311–328.

[R] G.A. Robertson: Strongly Chebyshev subspaces of matrices, *J. Approx. Theory* **55** (1988), 264–269.

[S] I. Singer, *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces*, Editura Academiei, Bucharest and Springer-Verlag, Berlin-Heidelberg-New York, 1970.

Addresses:

D. Hadwin: Department of Mathematics and Statistics, University of New Hampshire, Durham, NH 03824. E-mail: don@math.unh.edu.

D.R. Larson: Department of Mathematics, Texas A&M University, College Station, TX 77845. E-mail: larson@math.tamu.edu.

D. Timotin: Institute of Mathematics of the Romanian Academy, P.O. Box 1-764, Bucharest 014700, Romania. E-mail: Dan.Timotin@imar.ro.