J. OPERATOR THEORY
10(1983), 65–75

ON THE DISTANCE BETWEEN SIMILARITY ORBITS

DOMINGO A. HERRERO

1. INTRODUCTION

Consider the finite dimensional vector space $\mathbf{C}^n$ ($n \geqslant 1$) with its usual inner product and Hilbert space norm and let $\mathbf{M}_n(\mathbf{C})$ denote the Banach algebra of all $n \times n$ complex matrices under the norm $\|A\| = \max\{\|Ax\| : x \in \mathbf{C}^n, \|x\| = 1\}$. If $A, B \in \mathbf{M}_n(\mathbf{C})$, then a straightforward computation shows that

$$\|A' - B'\| \geqslant \frac{|\tau(A' - B')|}{n} = \frac{|\tau(A - B)|}{n} \tag{1}$$

for all $A', B' \in \mathbf{M}_n(\mathbf{C})$ such that A' is similar to A and B' is similar to B, where $\tau(R)$ denotes the *trace* of $R \in \mathbf{M}_n(\mathbf{C})$.

The main result of this note says that if A is a cyclic operator (this is equivalent to saying that the minimal monic polynomial of A coincides with $\mathrm{d}_A(\lambda) = \mathrm{determinant}(\lambda I - A)$) and B is not a multiple of the identity, then the above lower bound cannot be improved. More precisely, if for $T \in \mathbf{M}_n(\mathbf{C})$,

$$\mathscr{S}(T) = \{WTW^{-1} : W \in \mathbf{M}_n(\mathbf{C}) \text{ is invertible}\}$$

denotes the *similarity orbit* of T, then we have the following

THEOREM 1. *If $A, B \in \mathbf{M}_n(\mathbf{C})$ ($n \geqslant 2$), A is cyclic and B is not a multiple of the identity, then*

$$\mathrm{dist}[\mathscr{S}(A), \mathscr{S}(B)] \overset{\text{(def)}}{=} \inf\{\|A' - B'\| : A' \in \mathscr{S}(A),\ B' \in \mathscr{S}(B)\} = \frac{|\tau(A - B)|}{n}.$$

The case when $A = \lambda I$ for some complex λ will be treated separately (Theorem 8 below). An example will illustrate about the difficulties of the general case.

Several consequences can be derived from Theorem 1. Among others, we have

PROPOSITION 2. *If* $N \in \mathbf{M}_n(\mathbf{C})$ *is a normal operator such that* $1 \in \sigma(N)$ ($=$ *the spectrum of* N) *and* $\sigma(N) \subset \{\lambda \in \mathbf{C} : \operatorname{Re}\lambda \geqslant 0\}$, *then*

$$\operatorname{dist}[N, \{Q \in \mathbf{M}_n(\mathbf{C}) : Q^n = 0\}] > \frac{1}{2\sqrt{n}}.$$

If A is a nilpotent (equivalently: $\sigma(A) = \{0\}$), $\sigma(B) = \{0,1\}$ and $\operatorname{rank} B = 1$, then the result of Theorem 1 follows from Proposition 2.35 of [4] (see also [2, Example 2.4]). If N is positive hermitian, then the result of Proposition 2 is Proposition 2.30 of the same reference. For future purposes, it will be convenient to introduce the notation $T \sim R$ to indicate that $T, R \in \mathbf{M}_n(\mathbf{C})$ are similar operators.

The main result of this note was developed during a short visit to the University of California at San Diego. The author is deeply indebted to Professors L. C. Chadwick and J. W. Helton for several helpful discussions.

2. THE MAIN RESULT

Let $\{e_j\}_{j=1}^n$ be the canonical ONB of $\mathbf{C}^n$ and let $T = (t_{ij})_{i,j=1}^n$ denote the matrix of $T \in \mathbf{M}_n(\mathbf{C})$ with respect to this basis. We shall need two auxiliary results.

LEMMA 3. *Let* $T \in \mathbf{M}_n(\mathbf{C})$ *be a nonconstant operator* (i.e., $T \neq \lambda I$ *for all* $\lambda \in \mathbf{C}$). *Given* $\varepsilon > 0$ *there exists* U_ε *unitary in* $\mathbf{M}_n(\mathbf{C})$ *such that* $R = U_\varepsilon T U_\varepsilon^* = (r_{ij})_{i,j=1}^n$ *satisfies*

(i) $r_{12}r_{23} \ldots r_{n-1,n} \neq 0$, *and*

(ii) $\max_i \left| r_{ii} - \dfrac{\tau(T)}{n} \right| < \varepsilon$.

Proof. Let $\alpha = \dfrac{\tau(T)}{n}$. By a well-known result (see, e.g., [3, § 56, Exercise 6(a)]), there exists U unitary such that all the diagonal elements of UTU^* are equal to α.

We can assume without loss of generality that $U = I$. Since $T \neq \alpha I$, it readily follows that $t_{ij} \neq 0$ for some (i,j), $1 \leqslant i,j \leqslant n$, $i \neq j$. Thus, replacing if necessary $\{e_j\}_{j=1}^n$ by $\{e_{\pi(j)}\}_{j=1}^n$ for some permutation π of $\{1,2,\ldots,n\}$, we can directly assume that $|t_{12}| = \delta > 0$.

Assume that $t_{12}t_{23}\ldots t_{s-1,s} \neq 0$, but $t_{s,s+1} = 0$ for some $s \leqslant n-1$. Let $V(\zeta) \in \mathbf{M}_n(\mathbf{C})$ be defined by the relations $V(\zeta)e_j = e_j$ for $j \neq 2$ or $s+1$, $V(\zeta)e_2 = \cos\zeta\, e_2 - \sin\zeta\, e_{s+1}$, $V(\zeta)e_{s+1} = \sin\zeta\, e_2 + \cos\zeta\, e_{s+1}$ $(\zeta \in \mathbf{C})$; then $V(0) = I$, $V(\pi/2)e_j = e_j$ for $j \neq 2$ or $s+1$, $V(\pi/2)e_2 = -e_{s+1}$ and $V(\pi/2)e_{s+1} = e_2$, and therefore $V(\zeta)TV(\zeta)^{-1} = (t_{ij}(\zeta))_{i,j=1}^n$, where $t_{jj} = \alpha$ for all $j \neq 2$ or $s+1$, $t_{12}(0) = t_{12} \neq 0$, $t_{22}(0) = \alpha$, $t_{s,s+1}(0) = 0$, $t_{s+1,s+1}(0) = \alpha$, $t_{12}(\pi/2) = 0$ and $t_{s,s+1}(\pi/2) = t_{12} \neq 0$.

Since the entries of $V(\zeta)TV(\zeta)^{-1}$ are *entire functions* of ζ, there exists ζ_s, $0 < \zeta_s < \varepsilon/2$, such that $t_{12}(\zeta_s) \neq 0$, $t_{23}(\zeta_s) \neq 0$, $t_{s,s+1}(\zeta_s) \neq 0$ and

$$\max\{|t_{22}(\zeta_s) - \alpha|, |t_{s+1,s+1}(\zeta_s) - \alpha|\} < \varepsilon/n .$$

Define

$$T' := V(\zeta_s)TV(\zeta_s)^{-1} = V(\zeta_s)TV(\zeta_s)^* =: (t'_{ij})_{i,j=1}^n .$$

It is easily seen that T' is unitarily equivalent to T, $t'_{12}t'_{23} \dots t'_{s-1,s}t'_{s,s+1} \neq 0$ and $\max_i |t'_{ii} - \alpha| < \varepsilon/n$.

Now the re ult follows by an obvious inductive argument. ▨

LEMMA 4. *Let*

$$T(\zeta_2, \zeta_3, \dots, \zeta_n) = \begin{bmatrix} 0 & -t_{12} & -t_{13} & \dots & -t_{1,n-1} & -t_{1n} \\ \zeta_2 & 0 & -t_{23} & \dots & -t_{2,n-1} & -t_{2n} \\ -\zeta_3 & 0 & 0 & \dots & -t_{3,n-1} & -t_{3n} \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot \\ (-1)^{n-1}\zeta_{n-1} & 0 & 0 & \dots & 0 & -t_{n-1,n} \\ (-1)^n\zeta_n & 0 & 0 & \dots & 0 & 0 \end{bmatrix},$$

where $t_{12}t_{23} \dots t_{n-1,n} \neq 0$. *Given a monic polynomial* p *of degree* n *of the form* $p(\lambda) := \lambda^n + a_{n-2}\lambda^{n-2} + \dots + a_1\lambda + a_0$ *there exists* $\zeta_2, \zeta_3, \dots, \zeta_n \in \mathbf{C}$ *such that* $\det[\lambda I - T(\zeta_2, \zeta_3, \dots, \zeta_n)] = p(\lambda)$.

Proof. By developing the determinant by the first column, we obtain

$$\det[\lambda I - T(\zeta_2, \zeta_3, \dots, \zeta_n)] = \lambda|A_{n-1}(\lambda)| + \zeta_2 \begin{vmatrix} t_{12} & * \\ 0 & A_{n-2}(\lambda) \end{vmatrix} + \zeta_3 \begin{vmatrix} B_3(\lambda) & * \\ 0 & A_{n-3}(\lambda) \end{vmatrix} + \dots$$

$$\dots + \zeta_k \begin{vmatrix} B_k(\lambda) & * \\ 0 & A_{n-k}(\lambda) \end{vmatrix} + \dots + \zeta_{n-1} \begin{vmatrix} B_{n-1}(\lambda) & * \\ 0 & \lambda \end{vmatrix} + \zeta_n|B_n(\lambda)| =$$

$$= \lambda^n + \zeta_2 t_{12}\lambda^{n-2} + \zeta_3|B_3(\lambda)|\lambda^{n-3} + \dots + \zeta_k|B_k(\lambda)|\lambda^{n-k} + \dots$$

$$\dots + \zeta_{n-1}|B_{n-1}(\lambda)| + \zeta_n|B_n(\lambda)|,$$

where $A_k(\lambda)$ is an upper triangular $k \times k$ matrix with diagonal entries equal to

λ (so that $\det A_k(\lambda) = |A_k(\lambda)| = \lambda^k$) and

$$B_k(\lambda) = \begin{bmatrix} t_{12} & t_{13} & t_{14} & \cdots & t_{1,k-1} & t_{1k} \\ \lambda & t_{23} & t_{24} & \cdots & t_{2,k-1} & t_{2k} \\ & \lambda & t_{34} & \cdots & t_{3,k-1} & t_{3k} \\ & & \cdot & \cdots & \cdot & \cdot \\ & & & \cdots & \cdot & \cdot \\ & & & \cdots & \cdot & \cdot \\ & 0 & & & & \\ & & & & t_{k-2,k-1} & t_{k-2,k} \\ & & & & \lambda & t_{k-1,k} \end{bmatrix}, \quad k = 2, 3, \ldots, n-1.$$

An inductive computation of the determinants $|B_k(\lambda)|$ indicates that

$$\det[\lambda I - T(\zeta_2, \zeta_3, \ldots, \zeta_n)] = \lambda^n + (t_{12}\zeta_2 + q_2)\lambda^{n-2} +$$

$$+ (t_{12}t_{23}\zeta_3 + q_3)\lambda^{n-3} + \ldots + (t_{12}t_{23} \ldots t_{k-1,k}\zeta_k + q_k)\lambda^{n-k} + \ldots$$

$$\ldots + (t_{12}t_{23} \ldots t_{n-1,n}\zeta_n),$$

where q_k is a homogeneous polynomial of degree k in the variables $\{t_{ij}\}_{1 \leqslant i < j \leqslant k}$ and $\zeta_{k+1}, \zeta_{k+2}, \ldots, \zeta_n$.

Since $t_{12} \neq 0$, $t_{12}t_{23} \neq 0, \ldots, t_{12}t_{23} \ldots t_{n-1,n} \neq 0$, we can inductively define $\zeta_n = a_0(t_{12}t_{23} \ldots t_{n-1,n})^{-1}$, $\zeta_{n-1} = (a_1 - q_{n-1})(t_{12}t_{23} \ldots t_{n-2,n-1})^{-1}, \ldots, \zeta_k = (a_k - q_{n-k})(t_{12}t_{23} \ldots t_{k-1,k})^{-1}, \ldots, \zeta_2 = (a_{n-2} - q_2)t_{12}^{-1}$.

It is completely apparent that, with this choice of the coefficients $\zeta_2, \zeta_3, \ldots, \zeta_n$, we shall have $\det[\lambda I - T(\zeta_2, \zeta_3, \ldots, \zeta_n)] = p(\lambda)$. ▨

Now we are in a position to prove the main result. Let A and B be as in Theorem 1 and let $\varepsilon > 0$ be given. By Lemma 3 there exists $B' \sim B$ such that $B' = (b_{ij})_{i,j=1}^n$ with $b_{12}b_{23} \ldots b_{n-1,n} \neq 0$ and $\max_i \left| b_{ii} - \frac{\tau(B)}{n} \right| < \varepsilon$.

On the other hand, by Lemma 4 there exists $T = T(\zeta_2, \zeta_3, \ldots, \zeta_n)$ with $t_{ij} = b_{ij}$ for all (i,j) such that $1 \leqslant i < j \leqslant n$ and $\det[\lambda I - T] = \det\left[\lambda I - A + \frac{\tau(A)}{n}\right]$. Define $A' = T + \frac{|\tau(A)|}{n} I$; then $\det(\lambda I - A') = d_A(\lambda)$ and therefore A' and A have the same spectrum and, moreover, for each point in the spectrum the corresponding spectral invariant subspaces have the same dimension. Since A is cyclic, it follows from [4, Corollary 2.2] that A' belongs to the norm-closure $\mathscr{S}(A)^-$ of the similarity orbit of A.

Given $r > 1$, let $R_r \in \mathbf{M}_n(\mathbf{C})$ be the invertible matrix defined by $R_r e_j = r^j e_j$, $j = 1, 2, \ldots, n$; then $A_r = R_r A' R_r^{-1} \sim A'$, $A_r \in \mathscr{S}(A)^-$, $B_r = R_r B' R_r^{-1} \sim B$ and

$$B_r - A_r =$$

$$=: \begin{bmatrix} b_{11} - \tau(A)/n & & & & & \\ (b_{21}+\zeta_2)/r & b_{22} - \tau(A)/n & & & 0 & \\ & b_{32}/r & b_{33} - \tau(A)/n & & & \\ \cdot & & \cdot \quad \cdot & & & \\ \cdot & & \cdot \; \cdot & & & \\ \cdot & & \cdot \; \cdot & & & \\ \cdot & & \cdot & b_{n-1,n-1} - \tau(A)/n & & \\ (b_{n,1}+(-1)^n\zeta_n)/r^{n-1} & \cdot & \cdot & \cdot \quad \cdot & b_{n,n-1}/r & b_{nn} - \tau(A)/n \end{bmatrix}$$

so that

$$\operatorname{dist}[\mathscr{S}(A), \mathscr{S}(B)] \leqslant \|A_r - B_r\| \leqslant$$

$$\leqslant \max_i \left| b_{ii} - \frac{|\tau(A)|}{n} \right| + \frac{n^2}{r} (\|A'\| + \|B'\|) < \frac{|\tau(A)|}{n} + 2\varepsilon$$

provided r is large enough.

Since ε can be chosen arbitrarily small, we conclude (by using (1)) that $\operatorname{dist}[\mathscr{S}(A), \mathscr{S}(B)] = \dfrac{|\tau(A)|}{n}$. ▨

REMARKS. (i) Ad hoc modifications of the proof of Lemma 3 show that, if $\mathscr{U}_n$ denotes the unitary group of $\mathbf{C}^n$, then $\mathscr{U}_n(T) = \{U \in \mathscr{U}_n$: all the entries of UTU^* are different from $0\}$ is an open dense subset of $\mathscr{U}_n$ and $\{UTU^*\colon U \in \mathscr{U}_n(T)\}$ is an open dense subset of the unitary orbit of T.

(ii) Similarly, if $\mathscr{H}$ is a complex separable infinite dimensional space with an ONB $\{e_j\}_{j=1}^{\infty}$ and T is a nonconstant (bounded linear) operator, then $\mathscr{U}(T) = \{U\colon U$ is a unitary operator and all the entries of the matrix of UTU^* with respect to the given ONB are different from $0\}$ is a G_δ-dense subset of the unitary group of $\mathscr{H}$.

(iii) Condition (ii) of Lemma 3 cannot be replaced by "$r_{ii} = \dfrac{\tau(T)}{n}$ for all $i = 1, 2, \ldots, n$". Indeed, if

$$T = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \in \mathbf{M}_3(\mathbf{C}),$$

then T is a nilpotent of order 2 and rank 1. Thus, if

$$R = \begin{bmatrix} 0 & r_{12} & r_{13} \\ r_{21} & 0 & r_{23} \\ r_{31} & r_{32} & 0 \end{bmatrix}$$

and $r_{12}r_{23} \neq 0$, then $\operatorname{rank} R \geqslant 2$ and therefore R cannot be similar to T. A fortiori, R cannot be unitarily equivalent to T.

(iv) Let A and B be as in Theorem 1. If $A' \sim A$, $B' \sim B$ and $\|A' - B'\| = \frac{|\tau(A - B)|}{n}$, hen (as in the proof of [4, Proposition 2.17(i)]) we can easily check that $B' = A' - \frac{\tau(A - B)}{n} I$. It readily follows that B is similar to a translation of A. (In particular, B is cyclic too.)

Conversely, if $B = WAW^{-1} + \lambda I$ for some $\lambda \in \mathbf{C}$ and some invertible $W \in \mathbf{M}_n(\mathbf{C})$, then the distance $\operatorname{dist}[\mathscr{S}(A), \mathscr{S}(B)] = \inf\{\|A' - B'\| : A' \sim A, B' \sim B\}$ is actually attained and equal to $\|WAW^{-1} - B\| = |\lambda|$. It is clear that in this case, for each $A' \sim A$, there exists $B' \sim B$ such that $\|A' - B'\| = \operatorname{dist}[\mathscr{S}(A), \mathscr{S}(B)]$, i.e., $\mathscr{S}(A)$ and $\mathscr{S}(B)$ are "parallel" orbits.

(v) In contrast with Theorem 1, the main result of [1] shows that if $\mathscr{H}$ is an infinite dimensional Hilbert space and A and B are operators acting on $\mathscr{H}$, then "in most cases" $\mathscr{S}(A)^- \cap \mathscr{S}(B)^-$ contains a large family of normal operators. In particular, $\operatorname{dist}[\mathscr{S}(A), \mathscr{S}(B)] = 0$.

(vi) Theorem 1 remains true if the operator norm in $\mathbf{M}_n(\mathbf{C})$ is replaced by any norm $\|\cdot\|'$ such that, if $Ae_j = \lambda_j e_j$, $j = 1, 2, \ldots, n$, then $\|A\|' = \max_j |\lambda_j|$.

3. SOME CONSEQUENCES OF THEOREM 1

The result of Theorem 1 can be used to compute the distance between many different kinds of pairs of similarity-invariant subsets of $\mathbf{M}_n(\mathbf{C})$. For example, we have the following two (very simple) corollaries. Their proofs are immediate consequences of the theorem.

Corollary 5. *Let Λ_1 and Λ_2 be two nonempty subsets of* $\mathbf{C}$ *and let* $\mathscr{S}(\Lambda_j) = \{T \in \mathbf{M}_n(\mathbf{C}) : \sigma(T) \subset \Lambda_j\}$ $(j = 1, 2)$; *then*

$$\operatorname{dist}[\mathscr{S}(\Lambda_1), \mathscr{S}(\Lambda_2)] = \inf\left\{(1/n)\left|\sum_{k=1}^{n} (\lambda_k - \mu_k)\right| : \lambda_k \in \Lambda_1,\ \mu_k \in \Lambda_2\right\}.$$

COROLLARY 6. *If* $\Gamma_1 = \{\lambda_1, \lambda_2, \ldots, \lambda_r\}$ *and* $\Gamma_2 = \{\mu_1, \mu_2, \ldots, \mu_s\}$ *are two subsets of* $\mathbf{C}$ *containing at most* n *points and* $\mathcal{S}\mathcal{E}(\Gamma_j) = \{T \in \mathbf{M}_n(\mathbf{C}) : \sigma(T) = \Gamma_j\}$ $(j = 1,2)$, *then*

$$\operatorname{dist}[\mathcal{S}\mathcal{E}(\Gamma_1), \mathcal{S}\mathcal{E}(\Gamma_2)] = \min\left\{(1/n)\left|\sum_{h=1}^{r} m_h\lambda_h - \sum_{k=1}^{s} n_k\mu_k\right| : m_1, m_2, \ldots, m_r, n_1, n_2, \ldots, n_s \text{ are positive integers such that } \sum_{h=1}^{r} m_h = \sum_{k=1}^{s} n_k = n\right\}.$$

Theorem 1 also has the following infinite dimensional extension.

PROPOSITION 7. *Let A and B be two compact operators acting on an infinite dimensional Hilbert space. Assume that either*

(1) $\sigma(A) = \sigma(B) = \{0\}$, *or*

(2) *A is not an algebraic operator and B has infinite rank*, *or*

(3) $B \neq 0$ *and the restriction of A to each spectral invariant subspace corresponding to a singleton $\{\lambda\}$ is a cyclic operator on this (necessarily finite dimensional) subspace, for each* $\lambda \in \sigma(A)\setminus\{0\}$; *then*

$$\operatorname{dist}[\mathcal{S}(A), \mathcal{S}(B)] = 0.$$

Proof. Given $\varepsilon > 0$, it follows from the characterization of the closure of the similarity orbit of a compact operator (see [4, Proposition 8.5 and 8.6]) that

(1) In the first case, $A \sim A_0$ and $B \sim B_0$, where $\|A_0\| < \varepsilon$ and $\|B_0\| < \varepsilon$.

(2) In the second case, there exist $A' \in \mathcal{S}(A)^-$ and $B' \in \mathcal{S}(B)^-$, and $n = n(\varepsilon, A, B)$ such that $A' = (A_1 \oplus A_2 \oplus \ldots \oplus A_n) \oplus A_0$, $B' = (B_1 \oplus B_2 \oplus \ldots \oplus B_n) \oplus B_0$, $\|A_0\| < \varepsilon$, $\|B_0\| < \varepsilon$, A_j and B_j act on the same subspace of finite dimension n, A_j is cyclic in this subspace, $B_j \neq 0$,

$$\frac{|\tau(A_j)|}{n} < \varepsilon \quad \text{and} \quad \frac{|\tau(B_j)|}{n} < \varepsilon, \quad \text{for all } j = 1, 2, \ldots, n.$$

(3) In the third case, there exist $A' \in \mathcal{S}(A)^-$ and $B' \in \mathcal{S}(B)^-$, and $n = n(\varepsilon, A, B)$ such that $A' = A'' \oplus A_0$, $B' = B'' \oplus B_0$, $\|A_0\| < \varepsilon$, $\|B_0\| < \varepsilon$, A'' and B'' act on the same subspace of finite dimension n, A'' is cyclic in this subspace, $B'' \neq 0$,

$$\frac{|\tau(A'')|}{n} < \varepsilon \quad \text{and} \quad \frac{|\tau(B'')|}{n} < \varepsilon.$$

In either case, it follows from Theorem 1 that there exist $A_\varepsilon = A'_\varepsilon \oplus A_0 \in \mathcal{S}(A)^-$ and $B_\varepsilon = B'_\varepsilon \oplus B_0 \in \mathcal{S}(B)^-$ such that A'_ε and B'_ε act on the same space and

$$\operatorname{dist}[\mathcal{S}(A'_\varepsilon), \mathcal{S}(B'_\varepsilon)] < 2\varepsilon.$$

A fortiori,

$$\operatorname{dist}[\mathscr{S}(A), \mathscr{S}(B)] \leqslant \operatorname{dist}[\mathscr{S}(A_\varepsilon), \mathscr{S}(B_\varepsilon)] \leqslant$$

$$\leqslant \max\{\operatorname{dist}[\mathscr{S}(A'_\varepsilon), \mathscr{S}(B'_\varepsilon)], \operatorname{dist}[\mathscr{S}(A_0), \mathscr{S}(B_0)]\} <$$

$$< \max\{2\varepsilon, \|A_0\| + \|B_0\|\} = 2\varepsilon .$$

Since ε can be chosen arbitrarily small, we are done. ▨

REMARK. Proposition 7 is essentially the best possible result that we can deduce from Theorem 1. Indeed, if $B = 1_1 \oplus 0$ is a rank one projection and $A = \lambda_m \oplus K$, where λ_m denotes λ acting on a space of dimension $m \geqslant 1$ and K is a compact quasinilpotent such that $K^n \neq 0$ for all $n = 1,2, \dots$, then we have

(1) If either $\lambda = 0$ (m arbitrary), or $\lambda \neq 0$ and $m = 1$, then $\operatorname{dist}[\mathscr{S}(A), \mathscr{S}(B)] = 0$ by Proposition 7 (third case).

(2) However, if $\lambda \neq 0$ and $m \geqslant 2$, then for each $A' \sim A$ and $B' \sim B$ we can find a unit vector $x' = x'(A', B') \in \operatorname{kernel}(A' - \lambda) \cap \operatorname{kernel} B'$, therefore

$$\operatorname{dist}[\mathscr{S}(A), \mathscr{S}(B)] = \inf\{\|A' - B'\| : A' \sim A, B' \sim B\} \geqslant$$

$$\geqslant \inf\{\|(A' - B')x'\| : A' \sim A, B' \sim B\} = |\lambda| > 0.$$

($\operatorname{dist}[\mathscr{S}(A), \mathscr{S}(B)]$ is actually equal to $|\lambda|$, in this case.)

Furthermore, exactly the same argument shows that

$$\operatorname{dist}[\mathscr{S}(\lambda_2 \oplus T), \mathscr{S}(1_1 \oplus 0)] \geqslant |\lambda|,$$

for any operator T, not necessarily compact or quasinilpotent!

4. DISTANCE FROM A MULTIPLE OF THE IDENTITY TO A SIMILARITY ORBIT

THEOREM 8. *If* $\lambda \in \mathbf{C}$ *and* $B \in \mathbf{M}_n(\mathbf{C})$, *then*

$$\operatorname{dist}[\lambda I, \mathscr{S}(B)] = \operatorname{dist}[\mathscr{S}(\lambda I), \mathscr{S}(B)] = \operatorname{sp}(\lambda - B) \overset{(\text{def})}{=} \max\{|\lambda - \mu| : \mu \in \sigma(B)\}.$$

(where $\operatorname{sp}(R)$ *denotes the spectral radius of the operator* R*).*

Proof. It is obvious that $\mathscr{S}(\lambda I) = \{\lambda I\}$. If $\mu \in \sigma(B)$, then there exists a unit vector $x \in \mathbf{C}^n$ such that $Bx = \mu x$; then

$$\|\lambda I - B\| \geqslant \|(\lambda I - B)x\| = \|(\lambda - \mu)x\| = |\lambda - \mu|,$$

whence we deduce that $\operatorname{dist}[\lambda I, \mathscr{S}(B)] \geqslant \operatorname{sp}(\lambda I - B)$.

On the other hand, a simple analysis of the Jordan form of a matrix indicates that given $\varepsilon > 0$ there exists $B_\varepsilon \sim B$ such that

$$\|\lambda I - B_\varepsilon\| < \mathrm{sp}(\lambda I - B_\varepsilon) + \varepsilon = \mathrm{sp}(\lambda I - B) + \varepsilon.$$

The proof of Theorem 8 is now complete. ▨

REMARKS. (i) The main result of [1] implies, in particular, that the formula $\mathrm{dist}[\lambda I, \mathscr{S}(B)] = \mathrm{sp}(\lambda I - B)$ also holds for B acting on a separable infinite dimensional Hilbert space. Indeed, the separability is irrelevant in this case.

(ii) Let λ and B be as in Theorem 8; then there exists $B' \sim B$ such that $\|\lambda - B'\| = \mathrm{sp}(\lambda I - B)$ if and only if each $\mu \in \sigma(B)$ satisfying $|\lambda - \mu| = \mathrm{sp}(\lambda I - B)$ is a *simple* zero of the minimal polynomial of B.

The following example illustrates about the difficulties involved with a possible general formula for $\mathrm{dist}[\mathscr{S}(A), \mathscr{S}(B)]$, $A, B \in \mathbf{M}_n(\mathbf{C})$.

EXAMPLE 9. Let $E, Q \in \mathbf{M}_n(\mathbf{C})$ be two operators such that $E^2 = E \neq 0$ and $Q^2 = 0$. Then we have

(i) If $\mathrm{rank}\, E \leqslant \mathrm{rank}\, Q$ ($\leqslant n/2$), then [4, Proposition 2.19] and its proof show that $\mathrm{dist}[\mathscr{S}(E), \mathscr{S}(Q)] = 1/2$, (independently of n !).

(ii) On the other hand, if $\mathrm{rank}\, E > \mathrm{rank}\, Q$, $E' \sim E$ and $Q' \sim Q$, then there exists a unit vector $x' = x'(E', Q')$ such that $E'x' = x'$, but $Q'x' = 0$, so that

$$\|E' - Q'\| \geqslant \|(E' - Q')x'\| = \|x'\| = 1.$$

Since E is similar to a non-zero orthogonal projection P and $Q \sim \varepsilon Q$ for all $\varepsilon > 0$ (use the Jordan form of Q), we conclude that $1 \leqslant \mathrm{dist}[\mathscr{S}(E), \mathscr{S}(Q)] \leqslant \inf_{\varepsilon > 0} \|P - \varepsilon Q\| = 1$, so that $\mathrm{dist}[\mathscr{S}(E), \mathscr{S}(Q)] = 1$. (Once again, the result is independent of n.)

5. APPROXIMATION OF NORMAL OPERATORS WITH POSITIVE REAL PART BY NILPOTENT OPERATORS

Let $N \in \mathbf{M}_n(\mathbf{C})$ be a normal operator with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ (counted with multiplicity) and assume that $\|N - Q\| \leqslant \varepsilon$ for some nilpotent operator Q. Since the norm of the resolvent $(\lambda - N)^{-1}$ ($\lambda \notin \sigma(N)$) is equal to $(\mathrm{dist}[\lambda, \sigma(N)])^{-1}$, it follows as in, e.g., [4, Proposition 2.30] that

$$\sigma(N)_\varepsilon = \{\lambda \in \mathbf{C} : \mathrm{dist}[\lambda, \sigma(N)] \leqslant \varepsilon\}$$

is a connected neighborhood of 0.

By using Theorem 1 and proceeding exactly as in the above reference, we see that if $\lambda_1 = 1$ and $\operatorname{Re}\lambda_j \geqslant 0$ for all $j = 1, 2, \ldots, n$, then

$$\varepsilon \geqslant \|N - Q\| \geqslant \frac{|\tau(N-Q)|}{n} = \frac{|\tau(N)|}{n} \geqslant \operatorname{Re}\frac{\tau(N)}{n} >$$

(2)

$$> \frac{1}{n}\{1 + (1 - 2\varepsilon) + (1 - 4\varepsilon) + \ldots + (1 - 2m\varepsilon)\} = \frac{1}{n}(m+1)(1 - m\varepsilon),$$

where m is the integral part of $(2\varepsilon)^{-1}$.

It readily follows that $\varepsilon > \frac{1}{4n\varepsilon}$, i.e., $\varepsilon > \frac{1}{2\sqrt{n}}$, whence we obtain Proposition 2.

Example 10. Assume that N is a normal operator as in Proposition 2 such that $\|N - Q\| \leqslant \varepsilon$ for some nilpotent Q and that the eigenvalues of N are "ε-dense" in $D^+ = \{\lambda \in \mathbf{C} : |\lambda| \leqslant 1, \operatorname{Re}\lambda \geqslant 0\}$, in the sense that $D^+ \subset \sigma(N)_\varepsilon$. Then $\varepsilon > \frac{1}{4}\left(\frac{5}{n}\right)^{1/3}$.

Proof. Let Δ denote the equilateral triangle with vertices $1 - 2\varepsilon$, $1 - 2(m+1)\varepsilon$ (where m is defined so that $1 - 2(2m+2)\varepsilon \geqslant 0 > 1 - 2(2m+3)\varepsilon$) and $[1 - 2(m+1)\varepsilon] + \sqrt{3}m\varepsilon i$ and let Γ be the net of vertices of equilateral triangles obtained by dividing the sides of Δ into $2m$ equal parts. Then an estimate similar to that of (2) shows that

$$\operatorname{Re}\tau(N) \geqslant \sum_{\lambda \in \Gamma} \operatorname{Re}\lambda - [(1 - 2\varepsilon) + (1 - 4\varepsilon) + \ldots + (1 - 2(2m+1)\varepsilon)] \geqslant$$

$$\geqslant 2\{4m\varepsilon + 2(4m-2)\varepsilon + 3(4m-4)\varepsilon + \ldots + (2m-2)6\varepsilon +$$

$$+ (2m-1)4\varepsilon + 2m \cdot 2\varepsilon\} - \{2\varepsilon + 4\varepsilon + 6\varepsilon + \ldots + 2(2m+1)\varepsilon\} >$$

$$> 5(m+1)^3\varepsilon > \frac{5}{64\varepsilon^2}$$

$(m \geqslant m_0)$. Proceeding as in the proof of Proposition 2, we see that

$$\varepsilon \geqslant \|N - Q\| > \frac{5}{64n\varepsilon^2},$$

and therefore $\varepsilon > \frac{1}{4}\left(\frac{5}{n}\right)^{1/3}$, (for all n large enough).

This result strongly contrasts with [4, Proposition 2.28] which exhibits a normal operator $L \in \mathbf{M}_n(\mathbf{C})$ such that $\|L\| = 1 \in \sigma(L)$ and $\|L - Q\| < 5\left(\frac{\pi}{n}\right)^{1/2}$ for a suitable nilpont Q. (This normal operator has 0 trace and its eigenvalues are "uniformly sparsed" through the whole unit disk.)

Proposition 2 and Example 10 provide a strong support to the following.

CONJECTURE ([4, Conjecture 2.29]). *There exists an absolute constant* $C > 0$ *such that*

$$\inf\{\|N - Q\| : Q^n = 0\} \geqslant C/\sqrt{n}$$

for all normal operators N in $\mathbf{M}_n(\mathbf{C})$ such that $\|N\| = 1$ $(n = 1, 2, \ldots)$.

REFERENCES

1. APOSTOL, C.; HERRERO, D. A.; VOICULESCU, D., The closure of the similarity orbit of a Hilbert space operator, *Bull. Amer. Math. Soc.*, **6**(1982), 421–426.
2. AUPETIT, B.; ZEMÁNEK, J., Local behavior of the spectral radius in Banach algebras, *J. London Math. Soc.*, (2), **23**(1981), 171–178.
3. HALMOS, P. R., *Finite dimensional vector spaces*, D. Van Nostrand, Princeton, New Jersey, 1958.
4. HERRERO, D. A., *Approximation of Hilbert space operators.* I, Pitman Books Ltd., Research Notes in Math., **72**, London-Boston-Melbourne, 1982.

DOMINGO A. HERRERO
Department of Mathematics,
Arizona State University,
Tempe, Arizona 85287,
U.S.A.

Received March 17, 1982.