J. OPERATOR THEORY
52(2004), 113–132

LINEAR SPANS OF UNITARY AND SIMILARITY ORBITS OF A HILBERT SPACE OPERATOR

KENNETH R. DAVIDSON and LAURENT W. MARCOUX

Communicated by Nikolai K. Nikolski

Abstract. In this note, we show that if a bounded linear operator T acting on an infinite dimensional, separable, complex Hilbert space $\mathcal{H}$ is not of the form scalar plus compact, then every bounded linear operator on $\mathcal{H}$ can be written as a linear combination of 14 or fewer operators unitarily equivalent to T, as a linear combination of 6 or fewer operators similar to T, and as a sum of 8 or fewer operators similar to T. When T is not polynomially compact, the set of all sums of 2 operators similar to T is dense in $\mathcal{B}(\mathcal{H})$, while if T is polynomially compact, but not of the form scalar plus compact, then the set of sums of 3 operators similar to T is dense in $\mathcal{B}(\mathcal{H})$.

Keywords: *Unitary orbit, similarity orbit, numerical range, Lie ideals, polynomially compact.*

MSC (2000): 47A58.

There is an extensive literature about results which express operators on Hilbert space or elements of various algebras as sums or products of nice operators, such as projections, idempotents, symmetries, diagonal operators, etc. ([12], [15], [16], [19], [20], [21], [23], [24], [25], [27], [30]). In particular, see the survey paper by Wu ([31]). The purpose of this paper is to contribute to this arcana by showing that there is actually nothing special about using a nice operator, and that in fact almost any operator will do with reasonably small universal bounds.

More precisely, we investigate whether every operator may be expressed as a sum or linear combination of operators which are similar or even unitarily equivalent to a single given operator T. An immediate obstruction arises if T has the form scalar plus compact, as the set of such operators is a similarity invariant closed subspace. Surprisingly, this is the *only* obstruction!

We will show that when T is not of the form scalar plus compact, every operator is the linear combination of 6 operators similar to T and the sum (without scalar multiples) of 8 such operators. When using operators unitarily equivalent to T, there is an obvious norm obstruction for sums, so we only consider linear combinations and show that 14 terms suffice; and in fact, 8 terms is sufficient if the operator is not normal plus compact with the essential spectrum contained in a line segment. Even though these results are relatively small, they are too large to expect them to be sharp, and lower bounds are very difficult to obtain. The best result shows that at least 5 terms are needed in general for linear combinations of similarity orbits. So 6 is an excellent bound.

In this regard, we do significantly better when we only require the norm density of these sums. Precisely, we show that if T is not polynomially compact, then the set of sums of only two operators similar to T is dense in the set of all bounded operators. We show that three terms are needed if T is quadratically compact. This is used to prove that for all polynomially compact operators T which are not scalar plus compact, the sums of three operators similar to T is dense in $\mathcal{B}(\mathcal{H})$.

1. NOTATION

Let $\mathcal{H}$ be a complex, separable, infinite dimensional Hilbert space. We denote by $\mathcal{B}(\mathcal{H})$ the algebra of bounded linear operators on $\mathcal{H}$, and by $\mathcal{K}(\mathcal{H})$ the ideal of compact operators. The canonical map from $\mathcal{B}(\mathcal{H})$ to the *Calkin algebra* $\mathcal{B}(\mathcal{H})/\mathcal{K}(\mathcal{H})$ is denoted by π. Given $T \in \mathcal{B}(\mathcal{H})$, the *unitary orbit* of T is $\mathcal{U}(T) = \{U^*TU : U \text{ is unitary}\}$. We write $T \simeq R$ if $R \in \mathcal{U}(T)$, and we say that T and R are *unitarily equivalent*. Two operators T and R are *similar* if there exists an invertible operator S such that $S^{-1}TS = R$, and the corresponding *similarity orbit* of T is $\mathcal{S}(T) = \{S^{-1}TS : S \text{ is invertible}\}$. More generally, for $\mathcal{M} \subseteq \mathcal{B}(\mathcal{H})$, we define the unitary orbit $\mathcal{U}(\mathcal{M}) = \{U^*MU : M \in \mathcal{M},\ U \text{ is unitary}\}$ and similarity orbit $\mathcal{S}(\mathcal{M}) = \{S^{-1}MS : M \in \mathcal{M},\ S \text{ is invertible}\}$. We say that $\mathcal{M}$ is *unitarily invariant* (respectively *similarity invariant*) if $\mathcal{U}(\mathcal{M}) = \mathcal{M}$ (respectively $\mathcal{S}(\mathcal{M}) = \mathcal{M}$). We also will say that T and R are *approximately unitarily equivalent* (write $T \simeq_{\mathrm{a}} R$) if there is a sequence of unitary operators U_n such that $R = \lim_n U_n T U_n^*$, or equivalently if $\overline{\mathcal{U}(T)} = \overline{\mathcal{U}(R)}$. The operator T is said to be *quasinilpotent* if its spectrum $\sigma(T) = \{0\}$, and T is called *polynomially compact* if $\pi(T)$ is algebraic, i.e. if $p(\pi(T)) = 0$ for some polynomial p.

It was shown by Fong, Miers and Sourour in [14] that if $\mathcal{L} \subseteq \mathcal{B}(\mathcal{H})$ is a linear manifold, then $\mathcal{L} = \mathcal{U}(\mathcal{L})$ if and only if $\mathcal{L} = \mathcal{S}(\mathcal{L})$ and that both of these notions are equivalent to $\mathcal{L}$ being a *Lie ideal*, i.e. if $L \in \mathcal{L}$ and $A \in \mathcal{B}(\mathcal{H})$, then $AL - LA \in \mathcal{L}$. Moreover, they independently obtained Topping's result ([28]) that any Lie ideal $\mathcal{L}$ of $\mathcal{B}(\mathcal{H})$ is either contained in the algebra $\mathcal{T} = \mathbb{C}I + \mathcal{K}(\mathcal{H}) = \{\lambda I + K : \lambda \in \mathbb{C}, K \in \mathcal{K}(\mathcal{H})\}$ or equals $\mathcal{B}(\mathcal{H})$. Operators in $\mathcal{T}$ have been referred to as *thin* in [11], but this nomenclature is not universal.

Given a subset $\mathcal{M}$ of $\mathcal{B}(\mathcal{H})$, let $\Sigma_n\mathcal{M}$ denote the set of operators which are the sum of n elements of $\mathcal{M}$; and let $\mathbb{C}\mathcal{M} = \bigcup_{\lambda \in \mathbb{C}} \lambda\mathcal{M}$. From the Fong, Miers and Sourour result, it follows that if T is an operator and $T \notin \mathcal{T}$, then $\bigcup_{n=1}^{\infty} \Sigma_n \mathbb{C}\mathcal{U}(T)$ is a unitarily invariant subspace of $\mathcal{B}(\mathcal{H})$ which is not contained in $\mathcal{T}$, and thus must coincide with $\mathcal{B}(\mathcal{H})$. In other words, every operator $X \in \mathcal{B}(\mathcal{H})$ can be written as a finite linear combination of operators T_k, each of which is unitarily equivalent to T. A similar result holds with unitary equivalence replaced by similarity.

We define

$$\begin{aligned}\beta_T(X) &= \min\{n \in \mathbb{N} : X \in \Sigma_n \mathbb{C}\mathcal{U}(T)\}\\ \beta_T &= \sup\{\beta_T(X) : X \in \mathcal{B}(\mathcal{H})\}\\ \mu_T(X) &= \min\{n \in \mathbb{N} : X \in \Sigma_n \mathbb{C}\mathcal{S}(T)\}\\ \mu_T &= \sup\{\mu_T(X) : X \in \mathcal{B}(\mathcal{H})\}\\ \sigma_T(X) &= \min\{n \in \mathbb{N} : X \in \Sigma_n \mathcal{S}(T)\}\\ \sigma_T &= \sup\{\sigma_T(X) : X \in \mathcal{B}(\mathcal{H})\}.\end{aligned}$$

The notation σ_T should not be confused with the spectrum $\sigma(T)$ of T. Evidently, $2 \leqslant \mu_T \leqslant \min\{\beta_T, \sigma_T\}$.

In [20], the second author raised the question: for which operators T is $\beta_T < \infty$? The purpose of this note is to answer this question by proving that if $T \notin \mathcal{T}$, then $\beta_T \leqslant 14$, $\mu_T \leqslant 6$ and $\sigma_T \leqslant 8$. There is also a topological version of these results. We write

$$\begin{aligned}\overline{\beta}_T(X) &= \min\{n \in \mathbb{N} : X \in \overline{\Sigma_n \mathbb{C}\mathcal{U}(T)}\}\\ \overline{\beta}_T &= \sup\{\overline{\beta}_T(X) : X \in \mathcal{B}(\mathcal{H})\}\\ \overline{\mu}_T(X) &= \min\{n \in \mathbb{N} : X \in \overline{\Sigma_n \mathbb{C}\mathcal{S}(T)}\}\\ \overline{\mu}_T &= \sup\{\overline{\mu}_T(X) : X \in \mathcal{B}(\mathcal{H})\}\\ \overline{\sigma}_T(X) &= \min\{n \in \mathbb{N} : X \in \overline{\Sigma_n \mathcal{S}(T)}\}\\ \overline{\sigma}_T &= \sup\{\overline{\sigma}_T(X) : X \in \mathcal{B}(\mathcal{H})\}.\end{aligned}$$

Again, $2 \leqslant \overline{\mu}_T \leqslant \min\{\overline{\beta}_T, \overline{\sigma}_T\}$. Also it is clear that $\overline{\beta}_T \leqslant \beta_T$, $\overline{\mu}_T \leqslant \mu_T$ and $\overline{\sigma}_T \leqslant \sigma_T$. We will prove that if $T \notin \mathcal{T}$, then $\overline{\sigma}_T \leqslant 3$ and in fact $\overline{\mu}_T = \overline{\sigma}_T = 2$ if T is not polynomially compact.

2. SPANS AND SUMS OF UNITARY AND SIMILARITY ORBITS

2.1. Given two operators A and B, we shall say that A is a *dilation* of B — or equivalently, B is a *compression* of A — if we can find operators Z_2, Z_3 and Z_4 so that $A \simeq \begin{bmatrix} B & Z_2 \\ Z_3 & Z_4 \end{bmatrix}$. Our estimate for β_T will rely on the following result of Fong and Wu ([16]). Recall that an operator $D \in \mathcal{B}(\mathcal{H})$ is said to be *diagonal* if there exists an orthonormal basis $\{e_n\}_{n=1}^\infty$ of $\mathcal{H}$ and a bounded sequence $\{d_n\}_{n=1}^\infty \subseteq \mathbb{C}$ such that $De_n = d_n e_n$, $n \geqslant 1$. We denote by $\mathcal{D}$ the set of diagonal operators.

2.2. THEOREM. (Fong-Wu) *Every operator $X \in \mathcal{B}(\mathcal{H})$ admits a diagonal dilation D with $\|D\| \leqslant 2\|X\|$.*

Indeed, the factor 2 may be replaced with $1 + \varepsilon$. Any contraction dilates to a unitary; and any unitary may be dilated to a normal operator of finite spectrum whose convex hull contains the unit circle.

2.3. Recall that the *numerical range* of an operator $T \in \mathcal{B}(\mathcal{H})$ is the set $W(T) = \{(Te, e) : e \in \mathcal{H}, \|e\| = 1\}$. The *essential numerical range* of T is $W_{\mathrm{e}}(T) = \bigcap_{K \in \mathcal{K}(\mathcal{H})} \overline{W(T+K)}$. It is well-known that $W(T)$ and $W_{\mathrm{e}}(T)$ are non-empty and convex, and that $W_{\mathrm{e}}(T)$ is compact and contains the spectrum of $\pi(T)$ in $\mathcal{B}(\mathcal{H})/\mathcal{K}(\mathcal{H})$. Moreover, $\lambda \in W_{\mathrm{e}}(T)$ if and only if there exists an orthonormal set $\{e_n\}_{n=1}^{\infty}$ such that $\lim_{n\to\infty} (Te_n, e_n) = \lambda$. One also defines the *numerical radius* as $w(T) = \sup\{|\lambda| : \lambda \in W(T)\}$ and the *essential numerical radius* as $w_{\mathrm{e}}(T) = \sup\{|\lambda| : \lambda \in W_{\mathrm{e}}(T)\}$. One has

$$\frac{1}{2}\|T\| \leqslant w(T) \leqslant \|T\| \quad \text{and} \quad \frac{1}{2}\|\pi(T)\| \leqslant w_{\mathrm{e}}(T) \leqslant \|\pi(T)\|.$$

For more information regarding the essential numerical range, we refer the reader to Fillmore, Stampfli and Williams ([13]). We shall require the following result due to Anderson and Stampfli ([1]).

2.4. Theorem. (Anderson-Stampfli) (i) *$T \in \mathcal{T}$ if and only if $W_{\mathrm{e}}(T) = \{\lambda\}$ for some complex number λ.*

(ii) *Let $T \in \mathcal{B}(\mathcal{H})$ and suppose that $\alpha, \beta \in W_{\mathrm{e}}(T)$. Then T can be represented as*

$$\begin{bmatrix} D_1 & 0 & * \\ 0 & D_2 & * \\ * & * & * \end{bmatrix}$$

on $\mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3$, where each $\mathcal{H}_k$ is isomorphic to $\mathcal{H}$, and where $D_1 = \operatorname{diag}\{\alpha_n\}_{n=1}^{\infty}$ and $D_2 = \operatorname{diag}\{\beta_n\}_{n=1}^{\infty}$ are diagonal operators with $\lim_{n\to\infty} \alpha_n = \alpha$ and $\lim_{n\to\infty} \beta_n = \beta$.

2.5. Lemma. *Suppose $T \in \mathcal{B}(\mathcal{H})$ and 0 lies in the interior of $W_{\mathrm{e}}(T)$. There exists $\delta > 0$ so that if $W \in \mathcal{B}(\mathcal{H})$ with $\|W\| \leqslant \delta$, then W admits T as a dilation; i.e. there exist operators W_2, W_3 and W_4 so that $T \simeq \begin{bmatrix} W & W_2 \\ W_3 & W_4 \end{bmatrix}$.*

Proof. Since 0 lies in the interior of $W_{\mathrm{e}}(T)$, there exists $\varepsilon > 0$ so that $\varepsilon \mathrm{i}^k$ are in $W_{\mathrm{e}}(T)$ for $1 \leqslant k \leqslant 4$. Let $\delta = \varepsilon/3$. An easy generalization of Theorem 2.4 shows that $\mathcal{H}$ can be decomposed as $\bigoplus_{k=1}^{5} \mathcal{H}_k$ so that with respect to this decomposition, T can be represented as

$$\begin{bmatrix} D_1 & & & & * \\ & D_2 & & & * \\ & & D_3 & & * \\ & & & D_4 & * \\ * & * & * & * & * \end{bmatrix}$$

where each $D_k = \operatorname{diag}\{\alpha_n(k)\}_{n=0}^{\infty}$ for some sequence $\{\alpha_n(k)\}_{n=0}^{\infty}$ satisfying $\lim_{n\to\infty} \alpha_n(k) = \varepsilon \mathrm{i}^k$, $1 \leqslant k \leqslant 4$. Without loss of generality, we may assume that $|\alpha_n(k) - \varepsilon \mathrm{i}^k| < \varepsilon/100$, $n \geqslant 0$. If we set $B_n = \operatorname{diag}\{\alpha_n(1), \alpha_n(2), \alpha_n(3), \alpha_n(4)\}$ for $n \geqslant 0$, then $\bigoplus_{k=1}^{4} D_k \simeq \bigoplus_{n=0}^{\infty} B_n$. Since the numerical range of a complex matrix is convex and must contain its spectrum, it follows that the ball of radius 2δ is contained in $W(B_n)$ for all $n \geqslant 0$.

Now, given $n \geqslant 0$ and $c_n \in \mathbb{C}$ with $|c_n| \leqslant 2\delta$, we can choose a basis $\{w_{4n+1}, w_{4n+2}, w_{4n+3}, w_{4n+4}\}$ for the space $\mathcal{J}_n$ upon which B_n acts so that

$$B_n \simeq \begin{bmatrix} c_n & B_{n2} \\ B_{n3} & B_{n4} \end{bmatrix} \begin{matrix} \operatorname{span}\{w_{4n+1}\} \\ \operatorname{span}\{w_{4n+2}, w_{4n+3}, w_{4n+4}\} \end{matrix}.$$

Let $\mathcal{N}_1 = \operatorname{span}\{w_{4n+1}\}_{n=0}^{\infty}$ and $\mathcal{N}_2 = \operatorname{span}\{w_{4n+2}, w_{4n+3}, w_{4n+4}\}_{n=0}^{\infty}$. On decomposing $\bigoplus\limits_{k=1}^{5} \mathcal{H}_k$ as $\mathcal{N}_1 \oplus (\mathcal{N}_2 \oplus \mathcal{H}_5)$, we have

$$T \simeq \begin{bmatrix} C & * \\ * & * \end{bmatrix},$$

where $C = \operatorname{diag}\{c_n\}_{n=0}^{\infty}$.

Let $W \in \mathcal{B}(\mathcal{H})$ with $\|W\| \leqslant \delta$. By Theorem 2.2, W can be dilated to a diagonal operator D with $\|D\| \leqslant 2\delta$. From above, any such D is a compression of T. Hence for any $W \in \mathcal{B}(\mathcal{H})$ with $\|W\| \leqslant \delta$, there exists an operator of the form $\begin{bmatrix} W & * \\ * & * \end{bmatrix}$ in $\mathcal{U}(T)$. ∎

We isolate two calculations which will be used repeatedly in the sequel.

2.6. LEMMA. (i) *Suppose* $T = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$. *Then* $2(A \oplus D) \in \Sigma_2\mathcal{U}(T) \subseteq \Sigma_2\mathcal{S}(T)$.

(ii) *Let* $f : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ *be an arbitrary function. Suppose that* $\mathcal{M}_1$ *and* $\mathcal{M}_2$ *are unitarily invariant subsets of* $\mathcal{B}(\mathcal{H}) \simeq \mathcal{B}(\mathcal{H} \oplus \mathcal{H})$ *such that* $\mathcal{M}_1$ *contains every operator of the form* $A \oplus f(A)$ *for* $A \in \mathcal{B}(\mathcal{H})$ *and* $\mathcal{M}_2$ *contains every operator of the form* $B \oplus 0$ *for* $B \in \mathcal{B}(\mathcal{H})$. *Then* $\mathcal{B}(\mathcal{H}) = \mathcal{M}_1 + \mathcal{M}_2$.

Proof. (i) Observe that $T = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \simeq T_1 := \begin{bmatrix} A & -B \\ -C & D \end{bmatrix}$, and therefore $2(A \oplus D) = T + T_1 \in \Sigma_2\mathcal{U}(T)$.

(ii) It is well known (see for example [16]) that every operator X admits an infinite tridiagonal form $X = [Y_{ij}]$ for $i, j \geqslant 1$ with $Y_{ij} = 0$ if $|i - j| > 1$ with respect to some decomposition $\mathcal{H} = \bigoplus\limits_{n=1}^{\infty} \mathcal{H}_n$ into finite dimensional blocks. Now define subspaces $\mathcal{K}_1 = \bigoplus\limits_{n=0}^{\infty} \mathcal{H}_{4n}$, $\mathcal{K}_2 = \bigoplus\limits_{n=0}^{\infty} (\mathcal{H}_{4n+1} \oplus \mathcal{H}_{4n+3})$ and $\mathcal{K}_3 = \bigoplus\limits_{n=0}^{\infty} \mathcal{H}_{4n+2}$. Then with respect to $\mathcal{H} = \mathcal{K}_1 \oplus \mathcal{K}_2 \oplus \mathcal{K}_3$, X admits a 3×3 tridiagonal form

$$X = \begin{bmatrix} X_{11} & X_{12} & 0 \\ X_{21} & X_{22} & X_{23} \\ 0 & X_{32} & X_{33} \end{bmatrix}.$$

By our assumption, there exists an operator $Z \in \mathcal{B}(\mathcal{H})$ so that

$$M_1 = \begin{bmatrix} X_{11} & X_{12} & 0 \\ X_{21} & X_{22} & 0 \\ 0 & 0 & Z \end{bmatrix} \in \mathcal{M}_1.$$

Also, $M_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & X_{23} \\ 0 & X_{32} & X_{33} - Z \end{bmatrix}$ lies in $\mathcal{M}_2$, and hence $X = M_1 + M_2$ belongs to $\mathcal{M}_1 + \mathcal{M}_2$. ∎

Observe that by setting f to be the zero function, we may also conclude that $\mathcal{B}(\mathcal{H}) = \Sigma_2 \mathcal{M}_2$.

2.7. Lemma. *Let $\lambda \in \mathbb{C}$, $B \in \mathcal{B}(\mathcal{H})$ and $A = B \oplus B^{(\infty)} \oplus -B^{(\infty)}$. If $T \in \mathcal{B}(\mathcal{H})$ admits $\lambda I + A$ as a compression, then $B \oplus 0$ belongs to $\Sigma_4 \mathbb{C}\mathcal{U}(T)$.*

Proof. Suppose $T = \begin{bmatrix} \lambda I + A & Z_2 \\ Z_3 & Z_4 \end{bmatrix}$. By Lemma 2.6(i), $\Sigma_2 \mathcal{U}(T)$ contains

$$T_1 = 2\big((\lambda I + A) \oplus Z_4\big) = \big(2\lambda I + 2(B \oplus B^{(\infty)} \oplus -B^{(\infty)})\big) \oplus 2Z_4.$$

Since $-A \simeq B \oplus -B^{(\infty)} \oplus B^{(\infty)}$, we find that $\Sigma_2 \mathbb{C}\mathcal{U}(T)$ also contains

$$\begin{aligned} -T_1 &\simeq -2(\lambda I + A) \oplus -2Z_4 \\ &\simeq T_2 := \big(-2\lambda I + 2(B \oplus -B^{(\infty)} \oplus B^{(\infty)})\big) \oplus -2Z_4. \end{aligned}$$

Hence $T_1 + T_2 \simeq 4(B \oplus 0) \in \Sigma_4 \mathbb{C}\mathcal{U}(T)$, which is clearly sufficient. ∎

2.8. Theorem. *Let $\mathcal{M} \subseteq \mathcal{B}(\mathcal{H})$ and suppose that $\mathcal{M}$ is unitarily invariant. Suppose furthermore that for each $A \in \mathcal{B}(\mathcal{H})$, there is an operator $L(A)$ in $\mathbb{C}\mathcal{M}$ which admits A as a compression. Then $\mathcal{B}(\mathcal{H}) = \Sigma_6 \mathbb{C}\mathcal{M}$. That is, every operator in $\mathcal{B}(\mathcal{H})$ can be expressed as a linear combination of 6 or fewer operators in $\mathcal{M}$.*

Proof. Our assumption is that if we are given any $A \in \mathcal{B}(\mathcal{H})$, we can find $L(A) \in \mathcal{M}$ so that $L(A) \simeq \begin{bmatrix} A & Z_2 \\ Z_3 & Z_4 \end{bmatrix}$. By Lemma 2.6(i), $L_2(A) = A \oplus Z_4$ lies in $\Sigma_2 \mathbb{C}\mathcal{M}$. If $B \in \mathcal{B}(\mathcal{H})$ is arbitrary, then with $A = B \oplus B^{(\infty)} \oplus -B^{(\infty)}$ and $\lambda = 0$, it follows from Lemma 2.7 that $B \oplus 0 \in \Sigma_4 \mathbb{C}\mathcal{U}(L(A)) \subseteq \Sigma_4 \mathbb{C}\mathcal{M}$.

We then apply Lemma 2.6(ii) with $\mathcal{M}_1 = \Sigma_2 \mathbb{C}\mathcal{M}$ and $\mathcal{M}_2 = \Sigma_4 \mathbb{C}\mathcal{M}$ to conclude that $\mathcal{B}(\mathcal{H}) = \mathcal{M}_1 + \mathcal{M}_2 = \Sigma_6 \mathbb{C}\mathcal{M}$. ∎

2.9. Theorem. *Let $T \in \mathcal{B}(\mathcal{H})$ be an operator which is not of the form scalar plus compact.*

(i) *If 0 lies in the interior of the essential numerical range $W_{\mathrm{e}}(T)$ of T, then every operator $X \in \mathcal{B}(\mathcal{H})$ can be expressed as a linear combination of 6 or fewer operators unitarily equivalent to T. Thus $\beta_T \leqslant 6$.*

(ii) *If the essential numerical range $W_{\mathrm{e}}(T)$ has non-empty interior, then 8 operators unitarily equivalent to T will suffice, and so $\beta_T \leqslant 8$.*

(iii) *In the remaining case, T is normal plus compact with essential spectrum contained in a line segment. Then every operator $X \in \mathcal{B}(\mathcal{H})$ can be written as a linear combination of 14 or fewer operators unitarily equivalent to T, i.e. $\beta_T \leqslant 14$.*

Proof. (i) This is an immediate consequence of Lemma 2.5 and Theorem 2.8 applied to $\mathcal{U}(T)$.

(ii) Suppose $W_{\mathrm{e}}(T)$ contains the point λ in its interior. By Lemma 2.5, there exists $\delta > 0$ so that if $\|A\| < \delta$, then $A + \lambda I$ admits T as a dilation. Suppose $B \in \mathcal{B}(\mathcal{H})$ is arbitrary and let $A = (\delta/2\|B\|)(B \oplus B^{(\infty)} \oplus -B^{(\infty)})$. Using

Lemma 2.7 and the fact that $\Sigma_4\mathbb{C}\mathcal{U}(T)$ is invariant under scalar multiplication, we find that $B \oplus 0 \in \Sigma_4\mathbb{C}\mathcal{U}(T)$. The comment following Lemma 2.6 implies that $\Sigma_8\mathbb{C}\mathcal{U}(T) = \mathcal{B}(\mathcal{H})$.

(iii) When $W_{\mathrm{e}}(T)$ does not have interior yet T is not of the form scalar plus compact, $W_{\mathrm{e}}(T)$ contains at least two points. As $W_{\mathrm{e}}(T)$ is convex and compact, it must in fact be a line segment of the form $[\alpha, \beta]$. First suppose that $W_{\mathrm{e}}(T)$ is contained in $\mathbb{R}^+$. If A is the image of $\pi(T)$ under a faithful $*$-representation on $\mathcal{K}$, it follows that $W(A) \subset \mathbb{R}^+$. That is, $(Ax, x) \geqslant 0$ for all $x \in \mathcal{K}$, whence A is positive. Therefore T is positive plus compact. In general, by translating T by a scalar and multiplying by a non-zero constant, we reduce to this case; and hence T is normal plus compact with essential spectrum contained in the line segment $W_{\mathrm{e}}(T)$.

By Theorem 2.4, we can decompose $\mathcal{H}$ as $\bigoplus\limits_{k=1}^{4} \mathcal{H}_k$ so that the corresponding operator matrix for T is:

$$T \simeq T_1 = \begin{bmatrix} D_1 & & & * \\ & D_2 & & * \\ & & D_3 & * \\ * & * & * & * \end{bmatrix},$$

where $D_1 = \alpha I + K_1$, $D_2 = \alpha I + K_2$, $D_3 = \beta I + K_3$ and K_k are compact diagonal operators for $1 \leqslant k \leqslant 3$.

Then T is unitarily equivalent to

$$T_2 = \begin{bmatrix} D_3 & 0 & 0 & * \\ 0 & D_1 & 0 & * \\ 0 & 0 & D_2 & * \\ * & * & * & * \end{bmatrix} \simeq T_3 = \begin{bmatrix} D_2 & 0 & 0 & * \\ 0 & D_3 & 0 & * \\ 0 & 0 & D_1 & * \\ * & * & * & * \end{bmatrix}.$$

Now $T_1 + \mathrm{i}T_2 \in \Sigma_2\mathbb{C}\mathcal{U}(T)$, and $\{\alpha + \mathrm{i}\beta, \alpha + \mathrm{i}\alpha, \beta + \mathrm{i}\alpha\} \subseteq W_{\mathrm{e}}(T_1 + \mathrm{i}T_2)$. Since these three points are never collinear in $\mathbb{C}$, the convex set $W_{\mathrm{e}}(T_1 + \mathrm{i}T_2)$ has interior. The proof of (ii) shows that for any $B \in \mathcal{B}(\mathcal{H})$, $B \oplus 0$ belongs to $\Sigma_4\mathbb{C}\mathcal{U}(T_1 + \mathrm{i}T_2) \subseteq \Sigma_8\mathbb{C}\mathcal{U}(T)$.

Now note that $\Sigma_3\mathbb{C}\mathcal{U}(T)$ contains the operator $R = (\beta - \alpha)^{-1} \sum\limits_{k=1}^{3} \omega^k T_k$, where $\omega = \mathrm{e}^{2\pi\mathrm{i}/3}$. We claim that 0 lies in the interior of $W_{\mathrm{e}}(R)$. Indeed,

$$R = \begin{bmatrix} \omega^2(I + J_1) & 0 & 0 & * \\ 0 & (I + J_2) & 0 & * \\ 0 & 0 & \omega(I + J_3) & * \\ * & * & * & * \end{bmatrix}$$

where J_1, J_2 and J_3 are compact. It now follows from [13] that $\{1, \omega, \omega^2\}$ is contained in $W_{\mathrm{e}}(R)$, and hence 0 lies in the interior of the convex set $W_{\mathrm{e}}(R)$.

Using Lemma 2.5 and Lemma 2.6(i), we conclude that for any operator $A \in \mathcal{B}(\mathcal{H})$ there exists $Z \in \mathcal{B}(\mathcal{H})$ so that $A \oplus Z \in \Sigma_2\mathbb{C}\mathcal{U}(R) \subseteq \Sigma_6\mathbb{C}\mathcal{U}(T)$. Finally an application of Lemma 2.6(ii) tells us that $\mathcal{B}(\mathcal{H}) = \Sigma_{14}\mathbb{C}\mathcal{U}(T)$, completing the proof. ∎

Recall that a *unilateral* (respectively *bilateral*) *weighted shift* on $\mathcal{H}$ is a map $V \in \mathcal{B}(\mathcal{H})$ so that $Ve_n = v_n e_{n+1}$ where $\{e_n\}_{n\in\mathbb{N}}$ (respectively $\{e_n\}_{n\in\mathbb{Z}}$) is an orthonormal basis for $\mathcal{H}$ and $\{v_n\}_{n\in\mathbb{N}}$ (respectively $\{v_n\}_{n\in\mathbb{Z}}$) is a bounded sequence of complex numbers. It is well-known that the spectrum and essential spectrum of a weighted shift have circular symmetry, i.e. if $z \in \sigma(V)$ (respectively $z \in \sigma(\pi(V))$), then $\lambda z \in \sigma(V)$ (respectively $\lambda z \in \sigma(\pi(V))$) for all $|\lambda| = 1$. Moreover, when V is not compact, the spectral radius of $\pi(V)$ is strictly positive. Since $W_{\mathrm{e}}(V)$ is convex and contains $\sigma(\pi(V))$, the above theorem implies:

2.10. Corollary. *Let V be a non-compact weighted unilateral or bilateral shift. Then $\mathcal{B}(\mathcal{H}) = \Sigma_6\mathbb{C}\mathcal{U}(V)$. Thus $\beta_V \leqslant 6$.*

This improves a result of [20], where it was shown that every $X \in \mathcal{B}(\mathcal{H})$ is a sum of 18 or fewer weighted shifts, not necessarily unitarily equivalent to each other. It would be interesting to know if there exists an operator T so that $\beta_T < 6$, and more generally, what the minimum possible value of β_T is as T ranges over $\mathcal{B}(\mathcal{H})$.

2.11. Corollary. *If T is not normal plus compact with essential spectrum contained in a line segment, then $\mathcal{B}(\mathcal{H}) = \Sigma_8\mathbb{C}\mathcal{U}(T)$. Thus $\beta_T \leqslant 8$.*

2.12. Remark. If $T \in \mathcal{T}$, then $\mathcal{U}(T) \subseteq \mathcal{T}$ and any linear combination of operators in $\mathcal{T}$ is again in $\mathcal{T}$. In general, however, we do not expect $\bigcup\limits_{n\geqslant 1} \Sigma_n\mathbb{C}\mathcal{U}(T) = \mathcal{T}$. $\mathcal{K}(\mathcal{H})$ contains many ideals, including amongst others the ideal of finite rank operators and each of the Schatten p-classes $\mathfrak{S}_p(\mathcal{H})$. (See, for example, [10] for the definition and properties of these ideals.) If $T = \lambda I + K$ and K lies in such an ideal, then any linear combination of operators unitarily equivalent to T is again of the form $\alpha I + L$ where L lies in that ideal.

We also cannot expect a uniform bound β (independent of T) when examining this problem on a finite dimensional space. There, rank plays a factor, so that if $P \in \mathbb{M}_n(\mathbb{C})$ is a rank one projection, then the identity operator $I_n \in \mathbb{M}_n(\mathbb{C})$ can be expressed as a linear combination of no fewer than n copies of P. We do, however, draw the reader's attention to the paper [19] of Laurie, Mathes and Radjavi which deals with the question of which $n \times n$ matrices can be expressed as a sum of k idempotent matrices. Also Wang and Wu ([30]) obtain interesting results on expressing matrices (and operators) as sums of nilpotents of order 2.

These remarks apply equally well to the case of spans of similarity orbits.

For spans of similarity orbits, we have the following result:

2.13. Theorem. *Let $T \in \mathcal{B}(\mathcal{H})$ and suppose T is not of the form scalar plus compact. Then every bounded linear operator on $\mathcal{H}$ can be written as a linear combination of 6 or fewer operators similar to T. That is, $\mathcal{B}(\mathcal{H}) = \Sigma_6\mathbb{C}\mathcal{S}(T)$, and so $\mu_T \leqslant 6$.*

Proof. By Theorem 5 of [1], given $A \in \mathcal{B}(\mathcal{H})$ we can find Q invertible so that

$$Q^{-1}TQ \simeq \begin{bmatrix} A & Z_2 \\ Z_3 & Z_4 \end{bmatrix}$$

for some Z_2, Z_3 and Z_4. Since $\mathcal{S}(T)$ is clearly unitarily invariant, we can once again apply Theorem 2.8 to obtain the conclusion. ∎

2.14. Remarks. The case of projections has been well studied. In particular, it has been shown ([23], [21]) that every self-adjoint operator is a real linear combination of 5 projections, and hence every operator is a linear combination of at most 10 projections. One can verify that their formulae yield projections of infinite rank and nullity, so that one obtains $\Sigma_{10}\mathbb{C}\mathcal{U}(I \oplus 0) = \mathcal{B}(\mathcal{H})$ or $\mu_{I\oplus 0} \leqslant 10$.

On the other hand, it has been shown (c.f. [31] and [25]) that the sum of four idempotents cannot equal $-I$. Conversely, it was shown by Pearcy and Topping ([24]) that every bounded linear operator X is a sum of 5 idempotent operators. An examination of their proof shows that all five idempotents that they use are similar to $T = I \oplus 0$. So $\sigma_{I\oplus 0} = 5$.

The next theorem can be viewed as a generalization of the Pearcy and Topping result to arbitrary non-thin operators. This aspect of using operators similar to a single operator was not part of the considerations of Pearcy and Topping, and arose merely because the similarity classes of an idempotent are rather limited, being a function only of rank and nullity.

They also show that every operator is the sum of five nilpotents of order 2. In this case however, even though all nilpotents of order 2 are similar to an operator of the form $\begin{bmatrix} 0 & X \\ 0 & 0 \end{bmatrix}$, there are many different similarity classes and their argument does not actually imply our result for any nilpotent of order 2. However a lower bound comes from a result of Wang and Wu ([30]) where it is shown that sums of four nilpotents of order 2 are always commutators; and thus the identity is not such a sum. Hence $\mu_Q = \sigma_Q \geqslant 5$ for any Q with $Q^2 = 0$. Theorem 2.13 shows that $\sigma_Q = \mu_Q \leqslant 6$ if Q is not compact. This shows that Theorem 2.13 is rather close to being sharp.

We now turn to the question of sums, rather than linear combinations of the similarity orbit. One would expect, a priori, that many more terms would be needed. However, the same techniques require only two extra terms.

2.15. Theorem. *Let $T \in \mathcal{B}(\mathcal{H})$ and suppose T is not of the form scalar plus compact. Then every bounded linear operator on $\mathcal{H}$ can be written as a sum of 8 or fewer operators similar to T. Hence $\mathcal{B}(\mathcal{H}) = \Sigma_8\mathcal{S}(T)$, or equivalently, $\sigma_T \leqslant 8$.*

Proof. The proof of Theorem 5 of [1] shows that there is an operator Z so that T is similar to an operator of the form

$$\begin{bmatrix} 0 & I & * \\ 0 & 0 & * \\ * & * & Z \end{bmatrix}.$$

Given $A \in \mathcal{B}(\mathcal{H})$,

$$\begin{bmatrix} I & 0 \\ -A & I \end{bmatrix} \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ A & I \end{bmatrix} = \begin{bmatrix} A & I \\ -A^2 & -A \end{bmatrix},$$

and so T is also similar to

$$T_1 = \begin{bmatrix} A & I & * \\ -A^2 & -A & * \\ * & * & Z \end{bmatrix}.$$

The key issue is that the operator Z is *independent* of A.

Let $U = I \oplus \omega I \oplus \omega^2 I$, where $\omega = \mathrm{e}^{2\pi \mathrm{i}/3}$ is the cube root of unity. Then

$$T_1 + UT_1U^* + U^*T_1U = 3(A \oplus -A \oplus Z) \in \Sigma_3\mathcal{S}(T).$$

In particular, given $B \in \mathcal{B}(\mathcal{H})$, we can choose $A = (B \oplus -B \oplus Z \oplus -Z)^{(\infty)}$. Note that

$$3(A \oplus -A \oplus Z) \simeq 3A \simeq 3(B \oplus A) \simeq 3(B \oplus -A).$$

Adding the last two terms, we find that $6B \oplus 0 \in \Sigma_6\mathcal{S}(T)$. Since B was arbitrary, $C \oplus 0 \in \Sigma_6\mathcal{S}(T)$ for all $C \in \mathcal{B}(\mathcal{H})$.

Moreover, by Lemma 2.6(i), we also find that $2\left(A \oplus \begin{bmatrix} -A & * \\ * & Z \end{bmatrix}\right)$ belongs to $\Sigma_2\mathcal{S}(T)$ for all $A \in \mathcal{B}(\mathcal{H})$. So by Lemma 2.6(ii), it follows that $\mathcal{B}(\mathcal{H}) = \Sigma_8\mathcal{S}(T)$. ∎

We point out two special cases of interest where we can improve the estimate.

2.16. Proposition. *Let $Z \in \mathcal{B}(\mathcal{H})$. If T is similar to an operator of the form $\begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \oplus Z$ or $\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \oplus Z$, then $\Sigma_6\mathcal{S}(T) = \mathcal{B}(\mathcal{H})$, and hence $\sigma_T \leqslant 6$.*

Proof. Suppose first that T is similar to $\begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \oplus Z$. The proof is an easy modification of that of Theorem 2.15, if we simply note that given $A \in \mathcal{B}(\mathcal{H})$, T is similar to $\begin{bmatrix} A & I \\ -A^2 & -A \end{bmatrix} \oplus Z$ and so $2(A \oplus -A \oplus Z) \in \Sigma_2\mathcal{S}(T)$, from which the above calculation yields $C \oplus 0 \in \Sigma_4\mathcal{S}(T)$ for all $C \in \mathcal{B}(\mathcal{H})$, which in turn accounts for the reduction of 2 in the estimate for σ_T.

Next suppose that T is similar to $\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \oplus Z$. Then for any $A \in \mathcal{B}(\mathcal{H})$, $T \sim \begin{bmatrix} A & A - A^2 \\ I & I - A \end{bmatrix} \oplus Z$, and hence $2(A \oplus (I - A) \oplus Z) \in \Sigma_2\mathcal{S}(T)$. Fix $B \in \mathcal{B}(\mathcal{H})$ and let $A = (B \oplus (I - B) \oplus Z \oplus (I - Z))^{(\infty)}$. Then

$$A \oplus (I - A) \oplus Z \simeq A \simeq B \oplus A \simeq B \oplus (I - A).$$

Adding the last two expressions, we find that $4(B \oplus I) \in \Sigma_4\mathcal{S}(T)$.

Moreover, setting $f(A) = (I - A) \oplus Z$, we see that $2(A \oplus f(A)) \in \Sigma_2\mathcal{S}(T)$ for all $A \in \mathcal{B}(\mathcal{H})$. Given $X = \begin{bmatrix} X_{11} & X_{12} & 0 \\ X_{21} & X_{22} & X_{23} \\ 0 & X_{32} & X_{33} \end{bmatrix}$ arbitrary, it follows that

$$X = 2\begin{bmatrix} \frac{1}{2}(X_{11} - 4) & \frac{1}{2}X_{12} & 0 \\ \frac{1}{2}X_{21} & \frac{1}{2}X_{22} & 0 \\ 0 & 0 & Y \end{bmatrix} + 4\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & \frac{1}{4}X_{23} \\ 0 & \frac{1}{4}X_{32} & \frac{1}{4}(X_{33} - 2Y) \end{bmatrix}$$

lies in $\Sigma_6\mathcal{S}(T)$, where $Y = f\left(\begin{bmatrix} \frac{1}{2}(X_{11} - 4) & \frac{1}{2}X_{12} \\ \frac{1}{2}X_{21} & \frac{1}{2}X_{22} \end{bmatrix}\right)$. Thus $\sigma_T \leqslant 6$. ∎

We have already pointed out that lower bounds for these estimates are not easy to obtain. The following is an exception.

2.17. PROPOSITION. *Let $T \in \mathcal{B}(\mathcal{H})$. Not every operator can be expressed as a sum of two operators similar to T, and so $\sigma_T \geqslant 3$.*

Proof. Suppose otherwise, i.e. suppose $\sigma_T = 2$. Then for all $\lambda \in \mathbb{C}$, there exists $T_1, T_2 \in \mathcal{S}(T)$ so that $T_1 + T_2 = \lambda I$, and thus $T_1 = \lambda I - T_2$. But the spectra $\sigma(T_1)$ and $\sigma(T_2)$ of T_1 and T_2 coincide with $\sigma(T)$, whence $\sigma(T) = \lambda - \sigma(T)$ by the Spectral Mapping Theorem for every $\lambda \in \mathbb{C}$. This is evidently absurd. ∎

3. APPROXIMATION RESULTS

Suppose $T \notin \mathcal{T}$. The main result of this section is that if T is not polynomially compact, then $\overline{\mu}_T = 2$, and if T *is* polynomially compact, then $\overline{\mu}_T \leqslant 3$. Furthermore, we show that if T is quadratically compact, i.e. there is a polynomial p of degree 2 so that $p(T)$ is compact, then $\overline{\mu}_T = 3$. So 3 is in fact a sharp estimate in these cases.

3.1. Following [18], we shall write $A \underset{\text{sim}}{\longrightarrow} B$ to indicate that $B \in \overline{\mathcal{S}(A)}$. This is a transitive relation, so that if $A \underset{\text{sim}}{\longrightarrow} B$ and $B \underset{\text{sim}}{\longrightarrow} C$, then $A \underset{\text{sim}}{\longrightarrow} C$, but it is not symmetric.

Suppose that R is a universal quasinilpotent in the sense of [18], i.e. $\sigma(R) = \{0\}$ and $\pi(R^k) \neq 0$ for any $k \geqslant 1$. It was shown independently by Apostol ([2]) and Herrero ([17]) that the closure of the similarity orbit of R contains every other quasinilpotent operator (hence the terminology "*universal*"). More precisely, the closure of the similarity orbit of R coincides with the set of all biquasitriangular operators whose spectrum and essential spectrum are connected and contain the origin (c.f. [18]), where an operator $T \in \mathcal{B}(\mathcal{H})$ is *biquasitriangular* if both T and T^* have compact perturbations which are (upper) triangularizable with respect to some (potentially different) orthonormal bases ordered like $\mathbb{N}$. It is a deep result of Apostol, Foiaş and Voiculescu ([4]) that biquasitriangular operators are characterized by the index conditions $\mathrm{nul}(T - \lambda I) - \mathrm{nul}(T - \lambda I)^* = 0$ whenever $\pi(T - \lambda I)$ is either right or left invertible in the Calkin algebra.

3.2. REMARKS. We gather a number of facts which will be useful below.

(i) The following result is a special case of the Rosenblum-Davis-Rosenthal Corollary and can be found in [18], Corollary 3.22. The proof depends upon the invertibility of Rosenblum operators $\tau_{Z_{ii}, Z_{jj}}$ when the spectra of Z_{ii} and Z_{jj} are disjoint.

If $Z = \begin{bmatrix} Z_{11} & \cdots & \cdots & Z_{1n} \\ 0 & Z_{22} & \cdots & Z_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0. \cdots & 0 & Z_{nn} \end{bmatrix}$, and if $\sigma(Z_{ii}) \cap \sigma(Z_{jj}) = \emptyset$ if $i \neq j$, then Z is similar to $\bigoplus_{k=1}^{n} Z_{kk}$.

(ii) The *Apostol-Morrel simple models* ([5]; see also [18], Chapter 6.1) are bounded operators similar to a finite direct sum of essentially normal operators (i.e. $\pi(M_k)$ is normal in the Calkin algebra) with pairwise disjoint spectra. It was shown in [5] that every operator in $\mathcal{B}(\mathcal{H})$ can be approximated in norm by simple models.

Note that if M_k is essentially normal and $\beta \in \sigma_{\mathrm{e}}(M_k)$, then $M_k \simeq_{\mathrm{a}} M_k \oplus \beta I$ ([7]). From this it follows that the similarity orbit $\mathcal{S}(\mathcal{M})$ is dense in $\mathcal{B}(\mathcal{H})$ if $\mathcal{M}$ contains all operators of the form $A \oplus \lambda I$ where A is taken from any norm dense subset of $\mathcal{B}(\mathcal{H})$ and λ is restricted so that $\lambda \notin \sigma(A)$ or $0 < |\lambda| \leqslant \|A\|$. We will use these two variants. Thus to show that a similarity invariant set of operators is dense in $\mathcal{B}(\mathcal{H})$, it suffices to show that it contains one of these sets of operators.

(iii) Let $X \in \mathcal{B}(\mathcal{H})$, and suppose that ρ is a faithful representation of $C^*(\pi(X))$. From Voiculescu's non-commutative Weyl-von Neumann Theorem ([29]), we know that $X \simeq_{\mathrm{a}} X \oplus (\rho(\pi(X)))^{(\infty)}$. As a simple consequence of this result, if X is any operator and $n \geqslant 1$ is a positive integer, then X is approximately unitarily equivalent to an operator of the form $X_1 \oplus X_2 \oplus \cdots \oplus X_n$.

(iv) Suppose that L_1 and L_2 are nilpotents of order two, that both have 0 as an infinite dimensional direct summand, and that neither L_1 nor L_2 is compact. By Proposition 8.5 of [18] (see also Section 8.4 of the same reference), $L_1 \underset{\mathrm{sim}}{\longrightarrow} L_2$ and $L_2 \underset{\mathrm{sim}}{\longrightarrow} L_1$, or in other words, $\overline{\mathcal{S}(L_1)} = \overline{\mathcal{S}(L_2)}$. (Herrero attributes this result to C. Apostol and D. Voiculescu in an unpublished manuscript.) Indeed, any non-compact nilpotent of order 2 contains the operator $W = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \oplus 0$ in the closure of its similarity orbit.

3.3. Lemma. *Suppose that $A \simeq B \oplus C$, where $C \simeq C^{(\infty)}$. To show that $\overline{\sigma}_A \leqslant 3$, it suffices to show that $\mathcal{S}(B) + \mathcal{S}(C) + \mathcal{S}(C)$ is dense in $\mathcal{B}(\mathcal{H})$.*

Proof. Suppose that we can do this. Then for any $Y \in \mathcal{B}(\mathcal{H})$, it follows from 3.2 (iii) that $Y \simeq_{\mathrm{a}} Y \oplus Z \oplus Z$ for some $Z \in \mathcal{B}(\mathcal{H})$. Then

$$A \simeq A_1 := B \oplus C \oplus C \simeq A_2 := C \oplus B \oplus C \simeq A_3 := C \oplus C \oplus B.$$

Find invertible operators R_k and S_k for $k = 1, 2, 3$ so that

$$\|(R_1^{-1} B R_1 + R_2^{-1} C R_2 + R_3^{-1} C R_3) - Y\| < \varepsilon,$$

and

$$\|(S_1^{-1} B S_1 + S_2^{-1} C S_2 + S_3^{-1} C S_3) - Z\| < \varepsilon.$$

It easily follows that with $G_1 = R_1 \oplus S_2 \oplus S_2$, $G_2 = R_2 \oplus S_1 \oplus S_3$ and $G_3 = R_3 \oplus S_3 \oplus S_1$ we have

$$\|(G_1^{-1} A_1 G_1 + G_2^{-1} A_2 G_2 + G_3^{-1} A_3 G_3) - (Y \oplus Z \oplus Z)\| < \varepsilon,$$

and hence $Y \in \overline{\Sigma_3 \mathcal{S}(A)}$; and thus $\overline{\sigma}_A \leqslant 3$. ∎

3.4. LEMMA.
(i) *If* $T \simeq_{\mathrm{a}} T_1 \oplus T_2 \oplus \cdots \oplus T_m$*, then* $\overline{\sigma}_T \leqslant \max\{\overline{\sigma}_{T_i} : 1 \leqslant i \leqslant m\}$.
(ii) *If* $T \simeq_{\mathrm{a}} T_1 \oplus T_2 \oplus \cdots \oplus T_m$*, then* $\overline{\sigma}_T \leqslant 2\min\{\overline{\sigma}_{T_i} : 1 \leqslant i \leqslant m\}$.
(iii) $\overline{\sigma}_{sT+tI} = \overline{\sigma}_T$ *for all scalars* $s, t \in \mathbb{C}$ *with* $s \neq 0$.
(iv) *If* $A \in \overline{\mathcal{S}(T)}$*, then* $\overline{\sigma}_T \leqslant \overline{\sigma}_A$.

Proof. Since T and $T_1 \oplus T_2 \oplus \cdots \oplus T_m$ have the same closed unitary orbit, we may assume that T actually equals this direct sum for the purpose of computing $\overline{\sigma}_T$. Let $\max\{\overline{\sigma}_{T_i} : 1 \leqslant i \leqslant m\} = r$. By 3.2 (iii) above, given X in $\mathcal{B}(\mathcal{H})$, we can find $X_k \in \mathcal{B}(\mathcal{H})$ for $1 \leqslant k \leqslant m$ so that $X \simeq_{\mathrm{a}} X_1 \oplus \cdots \oplus X_m$. Let $\varepsilon > 0$ and choose $S_j(k)$ invertible in $\mathcal{B}(\mathcal{H})$ for $1 \leqslant j \leqslant r$ so that $\Big\| \sum\limits_{j=1}^{r} S_j(k)^{-1} T_k S_j(k) - X_k \Big\| < \varepsilon$. Then

$$\Big\| \sum_{j=1}^{r} \Big(\bigoplus_{k=1}^{m} S_j(k) \Big)^{-1} T \Big(\bigoplus_{k=1}^{m} S_j(k) \Big) - \bigoplus_{k=1}^{m} X_k \Big\| < \varepsilon,$$

which is clearly sufficient.

For part (ii), we may collapse all but one summand into a single term, so that $T \simeq_{\mathrm{a}} T_1 \oplus T_2$. We may suppose that $r = \overline{\sigma}_{T_1}$ takes the minimum value of the $\overline{\sigma}_{T_i}$. Again write an arbitrary operator as $X \simeq_{\mathrm{a}} X_1 \oplus X_2$. Choose invertible operators R_j and S_j for $1 \leqslant j \leqslant r$ so that

$$\Big\| (X_1 - rT_2) - \sum_{j=1}^{r} R_j^{-1} T_1 R_j \Big\| < \varepsilon, \quad \Big\| (X_2 - rT_2) - \sum_{j=1}^{r} S_j^{-1} T_1 S_j \Big\| < \varepsilon.$$

Then

$$\sum_{j=1}^{r} \begin{bmatrix} R_j & 0 \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix} \begin{bmatrix} R_j & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} 0 & I \\ S_j & 0 \end{bmatrix}^{-1} \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix} \begin{bmatrix} 0 & I \\ S_j & 0 \end{bmatrix}$$

approximates $X_1 \oplus X_2$ to within ε.

(iii) Suppose that $\Sigma_r \mathcal{S}(T)$ is dense in $\mathcal{B}(\mathcal{H})$, and consider $sT + tI$ for $s \neq 0$. Given $X \in \mathcal{B}(\mathcal{H})$, set $Y = s^{-1}(X - rtI)$. Choose R_k invertible for $1 \leqslant k \leqslant r$ so that $\Big\| \sum\limits_{k=1}^{r} R_k^{-1} T R_k - Y \Big\| < |s|^{-1}\varepsilon$. Then

$$\Big\| \sum_{k=1}^{r} R_k^{-1}(sT + tI)R_k - X \Big\| = |s| \, \Big\| \sum_{k=1}^{r} R_k^{-1} T R_k - Y \Big\| < \varepsilon.$$

Clearly this process is reversible, so $\overline{\sigma}_{sT+tI} = \overline{\sigma}_T$.

(iv) is straightforward. ∎

3.5. THEOREM. *Suppose* $T_1, T_2 \in \mathcal{B}(\mathcal{H})$ *are not polynomially compact. Then* $\mathcal{S}(T_1) + \mathcal{S}(T_2)$ *is dense in* $\mathcal{B}(\mathcal{H})$*. In particular,* $\overline{\sigma}_{T_1} = 2$.

Proof. Since T_k is not polynomially compact, it follows from [3], Theorem 9.1 that $T_k \underset{\mathrm{sim}}{\longrightarrow} rD$, where $r = \|T_1\| + \|T_2\|$ and D is a normal operator whose spectrum is equal to the closed unit disk $\overline{\mathbb{D}} = \{z \in \mathbb{C} : |z| \leqslant 1\}$. Given $X \in \mathcal{B}(\mathcal{H})$, index considerations show that $X \oplus (r + \|X\|)D$ is biquasitriangular and has spectrum containing $r\overline{\mathbb{D}}$. From the Similarity Orbit Theorem ([3], Theorem 9.1), we then find that $rD \underset{\text{sim}}{\longrightarrow} X \oplus (r + \|X\|)D$, and likewise $rD \underset{\text{sim}}{\longrightarrow} 0 \oplus \big(\alpha I - (r + \|X\|)D\big)$ if $|\alpha| \leqslant \|X\|$.

Thus $X \oplus \alpha I \in \overline{\Sigma_2 \mathcal{S}(rD)} \subseteq \overline{\mathcal{S}(T_1) + \mathcal{S}(T_2)}$. By 3.2 (ii), $\mathcal{S}(T_1) + \mathcal{S}(T_2)$ is dense in $\mathcal{B}(\mathcal{H})$. Take $T_2 = T_1$ to obtain $\overline{\sigma}_{T_1} = 2$. ∎

3.6. Lemma. *Suppose that P is a projection of infinite rank and nullity. Let $C \in \mathcal{B}(\mathcal{H})$ with $1 \notin \sigma(C)$. Then $\Sigma_3 \mathcal{S}(P \oplus C)$ is dense in $\mathcal{B}(\mathcal{H})$, i.e. $\overline{\sigma}_{P\oplus C} \leqslant 3$. This includes the case when C is non-existent.*

Proof. Observe that $P \simeq P^{(\infty)}$ and so $P \oplus C \simeq P^{(\infty)} \oplus (P \oplus C \oplus 0)$. Let $D \simeq C \oplus 0$. We will apply Lemma 3.3 to $P \oplus (P \oplus D)$. That is, it suffices to show that $\mathcal{S}(P) + \mathcal{S}(P) + \mathcal{S}(P \oplus D)$ is dense in $\mathcal{B}(\mathcal{H})$. Moreover, using 3.2 (ii), it suffices to approximate operators Y of the form $A \oplus \lambda I$ where A comes from a dense set of operators and $\lambda \notin \sigma(A)$.

We may assume that $A + (\lambda - 3)I - D$ is not scalar plus compact, for if it is not, we merely add a small perturbation to A to make it so. By the Brown-Pearcy Theorem ([8]), we can find $W, X \in \mathcal{B}(\mathcal{H})$ so that

$$A + (\lambda - 3)I - D = [W, X] := WX - XW.$$

Let $V = A - WX - D$. Then

$$Z_1 := \begin{bmatrix} WX & W \\ X - XWX & I - XW \end{bmatrix} \quad \text{and} \quad Z_2 := \begin{bmatrix} V & V \\ I - V & I - V \end{bmatrix}$$

are idempotents of infinite rank and nullity, and hence are similar to P.

Since $\sigma(D) \cap \{1\} = \big(\sigma(C) \cup \{0\}\big) \cap \{1\} = \emptyset$, it follows from 3.2 (i) that

$$Z_3 := \begin{bmatrix} D & -W - V \\ 0 & I \end{bmatrix} \sim D \oplus I \simeq D \oplus P.$$

We now compute

$$\begin{aligned} &Z_1 + Z_2 + Z_3 \\ &\quad = \begin{bmatrix} WX & W \\ X - XWX & I - XW \end{bmatrix} + \begin{bmatrix} V & V \\ I - V & I - V \end{bmatrix} + \begin{bmatrix} D & -W - V \\ 0 & I \end{bmatrix} \\ &\quad = \begin{bmatrix} A & 0 \\ X - XWX + I - V & \lambda I \end{bmatrix} \sim A \oplus \lambda I, \end{aligned}$$

where the similarity follows from Rosenblum's Theorem. By Lemma 3.3, this is sufficient. ∎

3.7. Corollary. *Let N be a normal operator which is not scalar plus compact. Then $\Sigma_3 \mathcal{S}(N)$ is dense in $\mathcal{B}(\mathcal{H})$, i.e. $\overline{\sigma}_N \leqslant 3$.*

Proof. First suppose that $\sigma_{\mathrm{e}}(N)$ contains $\{0, 1\}$. Then by the Weyl-von Neumann-Berg-Sikonia Theorem ([6]), it follows that $N \simeq_{\mathrm{a}} P \oplus N$ where $P = 0 \oplus I$ is a projection of infinite rank and nullity.

Split $N \simeq N_1 \oplus N_2$ so that $N_1 - I$ and N_2 are invertible. Then

$$P \oplus N \simeq (P \oplus N_1) \oplus \big((I - P) \oplus N_2\big).$$

By Lemma 3.6, $\overline{\sigma}_{P\oplus N_1} \leqslant 3$. Likewise, $\overline{\sigma}_{P\oplus(I-N_2)} \leqslant 3$. Then by Lemma 3.4(iii), the same holds for $I - \big(P \oplus (I - N_2)\big) = (I - P) \oplus N_2$. Finally, by Lemma 3.4(i), it follows that

$$\overline{\sigma}_N = \overline{\sigma}_{P\oplus N} \leqslant \max\{\overline{\sigma}_{P\oplus N_1}, \overline{\sigma}_{P\oplus(I-N_2)}\} \leqslant 3.$$

To finish, note that $\sigma_{\mathrm{e}}(N)$ contains at least two points. Thus for some scalars $s, t \in \mathbb{C}$ with $s \neq 0$, $\sigma_{\mathrm{e}}(sN + tI)$ contains $\{0, 1\}$. By Lemma 3.4(iii), $\overline{\sigma}_N = \overline{\sigma}_{sN+tI} \leqslant 3$. ∎

3.8. LEMMA. *Let* $W = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \oplus 0$ *and* $A \in \mathcal{B}(\mathcal{H})$ *be arbitrary. Then* $\Sigma_3\mathcal{S}(W \oplus A)$ *is dense in* $\mathcal{B}(\mathcal{H})$*, so that* $\overline{\sigma}_{W\oplus A} \leqslant 3$.

Proof. Since $W \simeq W^{(\infty)}$, by considering Lemma 3.3, it suffices to show that $\mathcal{S}(W) + \mathcal{S}(W) + \mathcal{S}(W \oplus A)$ is dense in $\mathcal{B}(\mathcal{H})$.

Let $0 \neq \lambda \in \mathbb{C}$, and suppose that R is a universal quasinilpotent. Denote by R_λ the operator $R - \lambda I$. Let $0 < \varepsilon < 1$. Since the spectral radius of R is 0, Rota's Theorem ([26]) (alternatively, Corollary 3.35 of [18]) implies that we can find an invertible operator B so that $\|B^{-1}R_\lambda B + \lambda I\| < \varepsilon$. For any $\alpha > 0$, we always replace B by αB and B^{-1} by $\alpha^{-1}B^{-1}$, so that we can assume *a priori* that $\|B\| < \varepsilon/\|R_\lambda\|$ at a cost of having $\|B^{-1}\|$ very large.

Let $N = \begin{bmatrix} R_\lambda & R_\lambda B \\ -B^{-1}R_\lambda & -B^{-1}R_\lambda B \end{bmatrix} \oplus 0$. Now both N and W are nilpotents of order two, both have 0 as an infinite dimensional direct summand, and neither N nor W is compact. By 3.4 (iv), $W \underset{\text{sim}}{\longrightarrow} N$ and $N \underset{\text{sim}}{\longrightarrow} W$. Then

$$A \oplus N \simeq \left[\begin{array}{cc|c} A & 0 & 0 \\ 0 & R_\lambda & R_\lambda \\ \hline 0 & -B^{-1}R_\lambda & -B^{-1}R_\lambda B \end{array}\right] \oplus 0 \simeq \begin{bmatrix} X & Y & 0 \\ Z & -B^{-1}R_\lambda B & 0 \\ 0 & 0 & 0 \end{bmatrix} =: M,$$

where $X = \begin{bmatrix} A & 0 \\ 0 & R_\lambda \end{bmatrix}$, $Y = \begin{bmatrix} 0 \\ R_\lambda \end{bmatrix}$ and $Z = [\,0 \quad -B^{-1}R_\lambda\,]$.

Let S_i denote contractive invertible operators to be chosen later, and define two nilpotents of order two similar to N and W respectively by

$$N_1 = (S_1 \oplus I \oplus I)N(S_1^{-1} \oplus I \oplus I) = \begin{bmatrix} S_1R_\lambda S_1^{-1} & S_1R_\lambda B & 0 \\ -B^{-1}R_\lambda S_1^{-1} & -B^{-1}R_\lambda B & 0 \\ 0 & 0 & 0 \end{bmatrix};$$

and with $V = -ZS_2^{-1}S_1R_\lambda^{-1}B$, let

$$\begin{aligned} N_2 &= \begin{bmatrix} S_1R_\lambda^{-1}B & 0 & 0 \\ 0 & I & 0 \\ 0 & V & I \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ I & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} B^{-1}R_\lambda S_1^{-1} & 0 & 0 \\ 0 & I & 0 \\ 0 & -V & I \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 \\ B^{-1}R_\lambda S_1^{-1} & 0 & 0 \\ -ZS_2^{-1} & 0 & 0 \end{bmatrix}. \end{aligned}$$

Finally, define an operator similar to $N \oplus A$ by

$$\begin{aligned}(S_2 \oplus I \oplus I)M(S_2^{-1} \oplus I \oplus I) &= \begin{bmatrix} S_2XS_2^{-1} & S_2YB & 0 \\ ZS_2^{-1} & -B^{-1}R_\lambda B & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &\simeq \begin{bmatrix} S_2XS_2^{-1} & 0 & S_2YB \\ 0 & 0 & 0 \\ ZS_2^{-1} & 0 & -B^{-1}R_\lambda B \end{bmatrix} =: N_3.\end{aligned}$$

Therefore $N_1, N_2 \in \overline{\mathcal{S}(W)}$ and $N_2 \in \overline{\mathcal{S}(W \oplus A)}$. Compute

$$\begin{aligned} &N_1 + N_2 + N_3 \\ &= \begin{bmatrix} S_1R_\lambda S_1^{-1} + S_2XS_2^{-1} & S_1R_\lambda B & S_2YB \\ 0 & -B^{-1}R_\lambda B & 0 \\ 0 & 0 & -B^{-1}R_\lambda B \end{bmatrix} \\ &= \begin{bmatrix} S_1R_\lambda S_1^{-1} + S_2XS_2^{-1} & 0 & 0 \\ 0 & \lambda I & 0 \\ 0 & 0 & \lambda I \end{bmatrix} - \begin{bmatrix} 0 & -S_1R_\lambda B & -S_2YB \\ 0 & B^{-1}R_\lambda B + \lambda I & 0 \\ 0 & 0 & B^{-1}R_\lambda B + \lambda I \end{bmatrix}.\end{aligned}$$

Consequently,

$$\big\|(N_1 + N_2 + N_3) - \big((S_1R_\lambda S_1^{-1} + S_2XS_2) \oplus \lambda I \oplus \lambda I\big)\big\| < 2\varepsilon.$$

As neither R_λ nor $X = A \oplus R_\lambda$ is polynomially compact, $\mathcal{S}(R_\lambda) + \mathcal{S}(A \oplus R_\lambda)$ is dense in $\mathcal{B}(\mathcal{H})$ by Theorem 3.5. From this we see that

$$\overline{\mathcal{S}(W) + \mathcal{S}(W) + \mathcal{S}(A \oplus W)} \ni T \oplus \lambda I$$

for any $T \in \mathcal{B}(\mathcal{H})$ and $\lambda \neq 0$. By 3.2 (ii), it follows that

$$\overline{\mathcal{S}(W) + \mathcal{S}(W) + \mathcal{S}(A \oplus W)} = \mathcal{B}(\mathcal{H}),$$

which concludes the proof. ∎

3.9. Lemma. *If A is a non-compact operator such that A^r is compact for some $r \geqslant 2$, then $A \underset{\mathrm{sim}}{\longrightarrow} A \oplus W$ where $W = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \oplus 0$.*

Proof. Let ρ denote a faithful $*$-representation of $C^*(\pi(A))$. Then $Q = \rho(\pi(A))^{(\infty)}$ is a non-compact nilpotent operator. By Voiculescu's non-commutative Weyl–von Neumann Theorem ([29]), $A \simeq_{\mathrm{a}} A \oplus Q$. Thus it suffices to show that $Q \underset{\mathrm{sim}}{\longrightarrow} W$. Define $\mathcal{H}_1 = \ker Q$, $\mathcal{H}_2 = \ker Q^2 \ominus \ker Q$ and $\mathcal{H}_3 = (\ker Q^2)^\perp$. With respect to the decomposition $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3$, we can write

$$Q = \begin{bmatrix} 0 & Q_{12} & Q_{13} \\ 0 & 0 & Q_{23} \\ 0 & 0 & Q_{33} \end{bmatrix}.$$

Clearly Q_{12} is injective, and Q_{33} is nilpotent. Since Q has infinite multiplicity, $Q_{12} \simeq Q_{12}^{(\infty)}$ is not compact.

Choose S_m so that $\|S_m^{-1}Q_{33}S_m\| < 1/m$ and $\|S_m\| < 1/m$. Then

$$\lim_{m\to\infty} (I \oplus I \oplus S_m)^{-1} Q (I \oplus I \oplus S_m) = \begin{bmatrix} 0 & Q_{12} \\ 0 & 0 \end{bmatrix} \oplus 0.$$

We can now appeal to 3.2 (iv) to conclude that $Q \underset{\mathrm{sim}}{\longrightarrow} W$. ∎

3.10. Theorem. *Suppose that T is not of the form scalar plus compact.*
(i) *If T is not polynomially compact, then $\Sigma_2\mathbb{C}\mathcal{S}(T)$ is dense in $\mathcal{B}(\mathcal{H})$, that is $\overline{\mu}_T = 2$.*
(ii) *If T is polynomially compact, then $\Sigma_3\mathbb{C}\mathcal{S}(T)$ is dense in $\mathcal{B}(\mathcal{H})$, that is $\overline{\mu}_T \leqslant 3$.*

Proof. Part (i) was proven in Theorem 3.5.

For part (ii), suppose that T is polynomially compact. Since $\pi(T)$ is algebraic, $\sigma_{\mathrm{e}}(T)$ is finite, say of cardinality n. We separate the argument into two cases.

If $n = 1$, then $\sigma_{\mathrm{e}}(T) = \{\lambda\}$ and so $T - \lambda I$ is essentially nilpotent but is not compact. By Lemma 3.9, $T \underset{\mathrm{sim}}{\longrightarrow} T \oplus (\lambda I + W)$. It follows by Lemma 3.4(iii) and (iv) and Lemma 3.8 that

$$\overline{\sigma}_T \leqslant \overline{\sigma}_{(\lambda I+W)\oplus T} = \overline{\sigma}_{W\oplus(T-\lambda I)} \leqslant 3.$$

If $n \geqslant 2$, apply Proposition 5.13 of [18] to observe that there is a normal operator N in $\overline{\mathcal{S}(T)}$ with $\sigma_{\mathrm{e}}(N) = \sigma_{\mathrm{e}}(T)$. So by Lemma 3.4(iv) and Corollary 3.7, $\overline{\sigma}_T \leqslant \overline{\sigma}_N \leqslant 3$. ∎

In general, the estimate of 3 for $\overline{\mu}_T$ obtained in part (ii) of the above theorem cannot be improved, as the following theorem and its corollary demonstrate.

3.11. Theorem. *Let $q(x) = (x-\alpha)(x-\beta)$ be a quadratic polynomial. Suppose that T_i are operators which are not scalar plus compact such that $q(T_i)$ is compact, $i = 1, 2$. Then*

$$\|T_1 + T_2 - \lambda I\| \geqslant \frac{|\alpha+\beta-\lambda|}{2}.$$

Therefore $\overline{\sigma}_{T_1} = 3$.

Proof. By Olsen's Theorem ([22]), there is a compact operator K_i so that $T_i - K_i \simeq \begin{bmatrix} \alpha I & R_i \\ 0 & \beta I \end{bmatrix}$. Hence it is a simple calculation to show that $W_{\mathrm{e}}(T_i)$ is an ellipse with foci α and β that depends only on $\|R_i\|$. Let r_i be the length of the major axis length of $W_{\mathrm{e}}(T_i)$, and let $\delta = |\alpha-\beta|/2$. Then the minor axis is $r_i' = (r_i^2 - \delta^2)^{1/2}$. Without loss of generality, $r_1 \geqslant r_2$, and thus $r_1' \geqslant r_2'$. It follows $W_{\mathrm{e}}(T_2)$ is contained in $W_{\mathrm{e}}(T_1)$.

Let $\varepsilon = |\alpha+\beta-\lambda|/2$, and suppose that $\|T_1 + T_2 - \lambda I\| < \varepsilon$. Then $T_1 = \lambda I - T_2 + X$ where $\|X\| < \varepsilon$. It follows that

$$d_H\big(W_{\mathrm{e}}(T_1), W_{\mathrm{e}}(\lambda I - T_2)\big) \leqslant \|X\| < \varepsilon$$

where d_H is the Hausdorff distance between compact sets. We claim that the Hausdorff distance between these two ellipses is at least ε, which will yield a contradiction.

Note that $W_{\mathrm{e}}(\lambda I - T_2) = \lambda - W_{\mathrm{e}}(T_2) = \mu + W_{\mathrm{e}}(T_2)$, where $\mu = (\lambda-\alpha-\beta)/2$, is the ellipse with foci $\lambda - \alpha$ and $\lambda - \beta$, centre $y = \lambda - x = x + \mu$, and major axis r_2. The supremum over the larger ellipse $W_{\mathrm{e}}(T_1)$ of the distance to the smaller one is at least as great as $d_H\big(W_{\mathrm{e}}(T_1), \mu + W_{\mathrm{e}}(T_1)\big) = |\mu| = \varepsilon$. To see that this latter distance is exactly ε, note that there is a certain point on the boundary of the first ellipse at which the tangent line is orthogonal to μ and the distance to the translated ellipse is exactly ε.

By Theorem 3.10, $\overline{\sigma}_T \leqslant 3$. However the argument above, using T_2 similar to T_1, shows that $\overline{\Sigma_2 \mathcal{S}(T_1)}$ contains only one scalar and thus $\overline{\sigma}_{T_1} \geqslant 3$. ∎

3.12. Corollary. *Suppose that $T \in \mathcal{B}(\mathcal{H})$ is not of the form scalar plus compact. If T^2 is compact, then:*

(i) $\overline{\Sigma_2 \mathbb{C}\mathcal{S}(T)} \neq \mathcal{B}(\mathcal{H})$*, and hence* $\overline{\mu}_T = 3$.

(ii) $\Sigma_4 \mathbb{C}\mathcal{S}(T) \neq \mathcal{B}(\mathcal{H})$*, and hence* $5 \leqslant \mu_T \leqslant 6$.

Proof. (i) The point here is that the set of essential nilpotents of order two is invariant under scalar multiplication as well as similarity. Thus $\overline{\mu}_T = \overline{\sigma}_T = 3$ by the previous theorem.

(ii) Any multiple of T is also an essential nilpotent of order 2. Suppose that I is the sum of four essential nilpotents of order two. Then in the Calkin algebra, the identity is the sum of four nilpotents of order two, and by taking a separable $*$-representation of the C^*-algebra that they generate, one obtains the identity as the sum of four nilpotents of order two. However, Wang and Wu ([30]) show that the sum of four nilpotents of order 2 is always a commutator; and the identity is not. So $\mu_T \geqslant 5$. On the other hand, by Theorem 2.8, $\mu_T \leqslant 6$. ∎

3.13. Example. Let T be a universal contraction

$$T := \bigoplus_{n=1}^{\infty} \bigoplus_{k=1}^{\infty} (T_{n,k} \oplus -T_{n,k})^{(\infty)}$$

where $\{T_{n,k}\}_{k=1}^{\infty}$ is a dense subset of the unit ball of $\mathbb{M}_n(\mathbb{C})$ for all $n \geqslant 1$. We take advantage of this opportunity to correct an error which occurred in the proof of the lower bound of 3 for $\overline{\beta}_T$ for this operator which appeared in Example 4.7 of [20]. In that paper, the second author correctly showed that $I \notin \overline{\Sigma_2 \mathbb{C}\mathcal{U}(T)}$, but then misstated the conclusion that $\overline{\beta}_T \geqslant 3$. Although the proof was incorrect, the result is nevertheless true. This can be seen as follows:

Suppose $\|I - \lambda_1 T_1 - \lambda_2 T_2\| < 1$ for some $\lambda_1, \lambda_2 \in \mathbb{C}$ and $T_1, T_2 \in \mathcal{U}(T)$. Since $T_k \simeq_{\mathrm{a}} \mathrm{e}^{\mathrm{i}\theta} T_k$ for all $\theta \in \mathbb{R}$, there is no loss of generality in assuming that each $\lambda_k \geqslant 0$, $k = 1, 2$, and that $\lambda_2 \leqslant \lambda_1$. Moreover, by replacing T_k by approximately unitarily equivalent copies of themselves, we may also assume that 1 and -1 are eigenvalues of T_k for $k = 1, 2$. Choose $x \in \mathcal{H}$ so that $\|x\| = 1$ and $T_1 x = -x$. Then from $\|(1 + \lambda_1)x - \lambda_2 T_2 x\| < 1$, and the fact that $\|T_2\| = 1$, we conclude that $\lambda_2 < \lambda_1$, a contradiction. Thus the operator I cannot be approximated by elements of $\Sigma_2 \mathbb{C}\mathcal{U}(T)$, and so $\overline{\beta}_T \geqslant 3$. As was shown there, $\overline{\beta}_T \leqslant 3$, and hence $\overline{\beta}_T = 3$.

On the other hand, by Theorem 3.10 , $\overline{\mu}_T = 2$.

Acknowledgements. Research supported in part by NSERC (Canada).

REFERENCES

1. J.H. Anderson, J.G. Stampfli, Commutators and compressions, *Israel J. Math.* **10**(1971), 433–441.
2. C. Apostol, Universal quasinilpotent operators, *Rev. Roumaine Math. Pures Appl.* **25**(1980), 135–138.
3. C. Apostol, L.A. Fialkow, D.A. Herrero, D. Voiculescu, *Approximation of Hilbert Space Operators.* II, Res. Notes in Math., vol. 102, Pitman Advanced Publishing Program, Boston–London–Melbourne, 1984.
4. C. Apostol, C. Foiaş, D. Voiculescu, Some results on nonquasitriangular operators. VI, *Rev. Roumaine Math. Pures Appl.* **18**(1973), 1473–1494.
5. C. Apostol, B.B. Morrel, On uniform approximation of operators by simple models, *Indiana Univ. Math. J.* **26**(1977), 427–442.
6. I.D. Berg, An extension of the Weyl-von Neumann theorem to normal operators, *Trans. Amer. Math. Soc.* **160**(1971), 365–371.
7. L. Brown, P.A. Fillmore, R.G. Douglas, Unitary equivalence modulo the compact operators and extensions of C^*-algebras, in *Proceedings of a Conference on Operator Theory, Halifax, Nova Scotia, 1973*, Lect. Notes in Math., vol. 345, Springer–Verlag, Berlin–Heidelberg–New York, 1973.
8. A. Brown, C. Pearcy, Structure of commutators, *Ann. of Math.* **82**(1965), 112–127.
9. K.R. Davidson, The distance between unitary orbits of normal operators, *Acta Sci. Math. (Szeged)* **50**(1986), 213–223.
10. K.R. Davidson, *Nest Algebras. Triangular Forms for Operators on a Hilbert Space*, Pitman Res. Notes Math. Ser., vol. 191, Longman Sci. Tech., Harlow 1988.
11. R.G. Douglas, C.M. Pearcy, A characterization of thin operators, *Acta Sci. Math. (Szeged)* **29**(1968), 295–297.
12. P.A. Fillmore, On products of symmetries, *Canad. J. Math.* **18**(1966), 897–900.
13. P.A. Fillmore, J.G. Stampfli, J.P. Williams, On the essential numerical range, the essential spectrum, and a problem of Halmos, *Acta Sci. Math. (Szeged)* **33**(1972), 179–192.
14. C.K. Fong, C.R. Miers, A.R. Sourour, Lie and Jordan ideals of operators on Hilbert space, *Proc. Amer. Math. Soc.* **84**(1982), 516–520.
15. C.K. Fong, A.R. Sourour, Sums and products of quasinilpotent operators, *Proc. Roy. Soc. Edinburgh Sect. A* **99**(1984), 193–200.
16. C.K. Fong, P.Y. Wu, Diagonal operators: dilation, sum and product, *Acta Sci. Math. (Szeged)* 57(1993), 125–138.
17. D.A. Herrero, Almost every quasinilpotent Hilbert space operator is a universal quasinilpotent, *Proc. Amer. Math. Soc.* **71**(1978), 212–216.
18. D.A. Herrero, *Approximation of Hilbert Space Operators.* I, Pitman Res. Notes Math. Ser., vol. 224, Longman Sci. Tech., Harlow, New York, second edition, 1989.
19. C. Laurie, B. Mathes, H. Radjavi, Sums of idempotents, *Linear Algebra Appl.* **208/209** (1994), 175–197.
20. L.W. Marcoux, Sums and products of weighted shifts, *Canad. Math. Bull.* **44** (2001), 469–481.
21. K. Matsumoto, Selfadjoint operators as a real span of 5 projections, *Math. Japon.* **29**(1984), 291–294.
22. C.L. Olsen, A structure theorem for polynomially compact operators, *Amer. J. Math.* **93**(1971), 686–698.
23. A. Paszkiewicz, Any self-adjoint operator is a finite linear combination of projectors, *Acad. Polon. Sci. Math.* **28**(1980), 227-245.

24. C. Pearcy, D. Topping, Sums of small numbers of idempotents, *Michigan Math. J.* **14**(1967), 453–465.
25. V.I. Rabanovich, Y.S. Samoĭlenko, When the sum of idempotents or projectors is a multiple of unity [Russian], *Funktsional. Anal. i Prilozhen.* **34**(2000), 91–93; translation in *Funct. Anal. Appl.* **34**(2000), 311–313.
26. G.-C. Rota, On models for linear operators, *Comm. Pure Appl. Math.* **13**(1960), 469–472.
27. J.G. Stampfli, Sums of projections, *Duke Math J.* **31**(1964), 455–461.
28. D. Topping, The unitarily invariant subspaces of $\mathcal{B}(\mathcal{H})$, preprint, 1970.
29. D. Voiculescu, A non-commutative Weyl-von Neumann theorem, *Rev. Roumaine Math. Pures Appl.* **21**(1976), 97–113.
30. J.H. Wang, P.Y. Wu, Sums of square-zero operators, *Studia Math.* **99**(1991), 115–127.
31. P.Y. Wu, Additive combinations of special operators, *Functional Analysis and Operator Theory (Warsaw, 1992)*, Banach Center Publ., vol. 30, Polish Acad. Sci., Warsaw 1994, pp. 337–361.

K.R. DAVIDSON
Department of Pure Mathematics
University of Waterloo
Waterloo, Ontario N2L 3G1
CANADA

E-mail: krdavids@math.uwaterloo.ca

L.W. MARCOUX
Department of Pure Mathematics
University of Waterloo
Waterloo, Ontario N2L 3G1
CANADA

E-mail: LWMarcoux@math.uwaterloo.ca

Received July 10, 2002.