THE SEMIGROUP GENERATED BY A SIMILARITY ORBIT OR A UNITARY ORBIT OF AN OPERATOR

C. K. FONG AND A. R. SOUROUR

Abstract. Let \( T \) be an invertible operator that is not a scalar modulo the ideal of compact operators. We show that the multiplicative semigroup generated by the similarity orbit of \( T \) is the group of all invertible operators. If, in addition, \( T \) is a unitary operator, then the multiplicative semigroup generated by the unitary orbit of \( T \) is the group of all unitary operators.

Introduction

Let \( H \) be a separable infinite-dimensional complex Hilbert space and let \( \mathcal{B}(H) \) be the algebra of all bounded operators on \( H \). We consider the following question: What is the multiplicative semigroup generated by the similarity orbit of an invertible operator on \( H \)? An analogous question for the unitary group is: What is the multiplicative semigroup generated by the unitary orbit of a unitary operator?

Let us call a subset \( S \) of a group \( G \) conjugation invariant, or simply invariant, if \( g^{-1}Sg \subseteq S \) for every \( g \in G \). (An invariant group is also called a normal subgroup.) One may ask what are the invariant semigroups of the group \( \mathcal{GL}(H) \) of invertible operators, or, respectively, of the group \( \mathcal{U}(H) \) of unitary operators.

We prove that if \( T \) is an invertible operator that is not a scalar modulo the ideal \( \mathcal{K}(H) \) of compact operators, then the multiplicative semigroup generated by the similarity orbit of \( T \) is the group of all invertible operators. Consequently, every proper invariant semigroup in \( \mathcal{GL}(H) \) is included in \( \mathcal{CI} + \mathcal{K}(H) \). This generalizes a theorem of Radjavi \([11]\) that asserts that every invertible operator is a product of a finite number (seven) of involutions, and a theorem of the authors \([7]\) that states that every invertible operator is a product of six unipotent operators.

Analogously, we show that if \( U \) is a unitary operator that is not a scalar modulo the compacts, then the semigroup generated by the unitary orbit of \( U \) is the group of all unitary operators. Consequently, every proper invariant semigroup in \( \mathcal{U}(H) \) is included in \( \mathcal{CI} + \mathcal{K}(H) \). This generalizes a theorem of Halmos and Kakutani \([9]\); namely, that every unitary operator is a product of four symmetries (i.e., self-adjoint unitary operators).

In the last section we prove a result about invariant groups in the Calkin algebra \( \mathcal{B}(H)/\mathcal{K}(H) \).

Received by the editors November 22, 2000 and, in revised form, May 17, 2002.

2000 Mathematics Subject Classification. Primary 47D03; Secondary 20F38.

Key words and phrases. Semigroups, conjugation-invariant.

This research was supported in part by an NSERC grant.
We end this introduction by noting that an additive version of the results in this paper is in [6]. A special case of the results in [6] is that every proper linear subspace of $B(H)$ that is invariant under conjugation by all invertible operators (respectively, all unitary operators) is included in $CI + K(H)$. We also note that semigroups generated by a similarity orbit of a matrix have been investigated in [8].

1. Statements of results

We start by stating the results about unitary operators.

**Theorem A.** Let $U$ be a unitary operator that is not the sum of a scalar and a compact operator. Then every unitary operator is a product of a finite number of operators each of which is unitarily equivalent to $U$.

The following is an immediate corollary.

**Corollary 1.** Every proper invariant semigroup (in particular, normal subgroup) in the group of unitary operators is included in $CI + K(H)$.

In Theorem A if we take $U$ to be a symmetry (i.e., a unitary operator $U$ satisfying $U^2 = I$), and if we also assume that both ker($U - I$) and ker($U + I$) are infinite dimensional, then we recover the qualitative part of the Halmos-Kakutani [9] result that states that every unitary operator is a product of four symmetries.

The theorem of Halmos and Kakutani has a “skew” version due to Radjavi [11]; namely, that every invertible operator is a product of seven involutions. (An *involution* is an operator whose square is the identity.) We also have the following “skew” version of Theorem A.

**Theorem B.** Let $T$ be an invertible operator which is not the sum of a scalar and a compact operator. Then every invertible operator is a product of a finite number of operators each of which is similar to $T$.

As before, we conclude the following about invariant semigroups.

**Corollary 2.** Every proper invariant semigroup (in particular, normal subgroup) in the group of invertible operators is included in $CI + K(H)$.

The following are special cases of Theorem B. First recall that an operator is said to be *unipotent* if it is the sum of the identity and a nilpotent operator, and is said to be a *unipotent of order* 2 if it is of the form $I + N$, where $N^2 = 0$.

**Corollary 3.** Every invertible operator is a product of a finite number of

(a) involutions (cf. [11]);
(b) unipotents of order 2 (cf. [7]);
(c) invertible positive operators (cf. [10]).

We again observe that in [11], [7] and [10], the number of factors are seven, six and seven, respectively. See also [13].

**Proof.** Parts (a) and (b) are obvious. To prove part (c), let $P$ be an invertible positive operator that is not a scalar plus compact. By Theorem B every invertible operator is a product of a finite number of operators each of which is similar to $P$. Each factor $S^{-1}PS$ is a product of two invertible positive operators since $S^{-1}PS = S^{-1}(S^{-1})^* (S^* PS)$. 

$\square$
We end this section with the following remarks about the number of factors in Theorems A and B.

**Remarks.** The number of factors in Theorem A is unbounded. Indeed, if $U$ is a unitary operator satisfying $\|U - I\| \leq 2^{-n}$, and if $V$ is a product of $n$ operators from the unitary orbit of $U$, then it is easy to see that $\|V - I\| \leq 1$. On the other hand, the proof of Theorem B given below establishes that 112 factors suffice for the factorization of that theorem. This is undoubtedly not a sharp estimate, but we make no attempt in the present work to investigate the minimum number of factors required.

2. **Proof of Theorem A**

We denote the essential numerical range of an operator $A$ by $W_e(A)$. For basic properties of the essential numerical range, the reader is referred to [5].

**Lemma 1.** If $U$ is a unitary operator and if zero is in the interior of the numerical range of $U$, then every unitary operator is a product of at most eight operators each of which is unitarily equivalent to $U$.

**Proof.** We denote the interior of the numerical range of $U$ by $W_e(U)^o$. Construct inductively an orthonormal sequence $\{e_n\}$ such that $(Ue_n, e_m) = 0$ for all $n, m$, as follows. Since $0 \in W_e(U)^o \subseteq W(U)$, there is a unit vector $e_1$ such that $(Ue_1, e_1) = 0$. Suppose now that we already have $e_1, \ldots, e_k$ such that $(Ue_n, e_m) = 0$ for all $n, m \leq k$. Let

$$M = \{e_1, \ldots, e_k, Ue_1, \ldots, Ue_k, U^*e_1, \ldots, U^*e_k\}^\perp$$

and let $V$ be the compression of $U$ to $M$. Since $M^\perp$ is finite dimensional, we have $W_e(V) = W_e(U)$ and hence $0 \in W_e(V)^o$. Let $e_{k+1}$ be a unit vector in $M$ such that $(Ve_{k+1}, e_{k+1}) = 0$. Then $e_1, \ldots, e_{k+1}$ is a finite orthonormal sequence such that $(Ue_n, e_m) = 0$ for all $n, m \leq k + 1$.

Let $H_1$ be the closed linear span of $\{e_n: n \text{ odd}\}$, let $H_3 = UH_1$ and let $H_2 = (H_1 \oplus H_3)^\perp$. The unitary operator $U$ maps $H_1$ onto $H_3$ and hence it maps $H_2 \oplus H_3 = H_1^\perp$ onto $H_3^\perp = H_1 \oplus H_2$ and so the matrix of $U$ relative to the decomposition $H = H_1 \oplus H_2 \oplus H_3$ takes the form

$$U = \begin{pmatrix} 0 & * & * \\ 0 & * & * \\ R & 0 & 0 \end{pmatrix}$$

where $R$ is a unitary operator from $H_1$ onto $H_3$. We note that each of $H_1$, $H_2$ and $H_3$ is isomorphic to $H$.

Now let $V$ be any unitary operator on $H_3$ and let

$$V_0 = \begin{pmatrix} 0 & 0 & R^*V \\ 0 & 1 & 0 \\ R & 0 & 0 \end{pmatrix}.$$ 

Then $V_0$ is a unitary operator on $H$ and

$$UV_0UV_0^* = \begin{pmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & V \end{pmatrix} = \begin{pmatrix} V' & 0 \\ 0 & V \end{pmatrix}.$$
Identifying each of $H_1 + H_2$ and $H_3$ with $H$, the above computation shows that if $V$ is a unitary operator on $H$, then there exists another unitary operator $V'$ such that $V \oplus V'$ is a product of two operators unitarily equivalent to $U$. We now take $V$ to be a bilateral shift of infinite multiplicity. The unitary operators $V'$ can be written as a product $V_1V_2$ of two bilateral shifts of infinite multiplicity [9]. Let $J$ be a unitary operator such that $V^* = JV_1J^*$ and let $S = JV_2J^*$. It follows that $V \oplus V'$ is unitarily equivalent to $V \oplus V^*S$ and so each of $V \oplus V^*S$ and $V^*S \oplus V$ is a product of two operators unitarily equivalent to $U$. So, there exist four operators unitarily equivalent to $U$ whose product is the operator $(V \oplus V^*S)(V^*S \oplus V) = S \oplus V^*SV$ which is a bilateral shift of infinite multiplicity. Now the conclusion of the lemma follows by using, once again, the fact that every unitary operator is a product of two bilateral shifts of infinite multiplicity. 

**Proof of Theorem** Suppose that $U$ is a unitary operator which is not a scalar plus compact. The essential spectrum $\sigma_e(U)$ of $U$ contains two distinct complex numbers $\lambda_1$ and $\lambda_2$. We may write $U$ in the form

$$U = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & A \end{pmatrix} + K_1$$

where $K_1$ is a compact operator and where every direct summand is infinite dimensional (see, e.g., [10, Theorem 4.2]). In view of Lemma 1 it suffices to show that there is a product $V$ of a finite number of operators unitarily equivalent to $U$ such that $0 \in W_c(V)$. 

We consider two cases according as $\lambda_2 = -\lambda_1$ or not. In the first case,

$$U = \lambda_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & B \end{pmatrix} + K_1$$

which is unitarily equivalent to

$$\lambda_1 \begin{pmatrix} 0 & J & 0 \\ J^* & 0 & 0 \\ 0 & 0 & B \end{pmatrix} + K_1$$

for every unitary operator $J$. Now let $R$ be a unitary operator such that $0 \in W_c(R)$. It follows that $U$ is unitarily equivalent to each of the operators

$$U_1 = \lambda_1 \begin{pmatrix} 0 & R & 0 \\ R^* & 0 & 0 \\ 0 & 0 & B \end{pmatrix} + K_1$$

and

$$U_2 = \lambda_1 \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & B \end{pmatrix} + K_1,$$

hence

$$U_1U_2 = \lambda_1^2 \begin{pmatrix} R & 0 & 0 \\ 0 & R^* & 0 \\ 0 & 0 & B^2 \end{pmatrix} + K_2$$

where $K_2$ is compact. Therefore $0 \in W_c(U_1U_2)$. This ends the proof in this case.
Finally, we consider the case \( \lambda_2 \neq -\lambda_1 \). Let \( \mu = \lambda_2 / \lambda_1 \), so \( \mu \neq \pm 1 \). It is easy to see that there exists a positive integer \( n \) such that 0 belongs to the interior of the convex hull of \( \{1, \mu, \mu^2, \ldots, \mu^n\} \). For every positive integer \( m \), we have

\[
U^m = \lambda_1^m \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mu^m & 0 \\ 0 & 0 & B^m \end{pmatrix} + K_m
\]

where \( K_m \) is compact. So \( U^m \) is unitarily equivalent to the operator

\[
V_m = \lambda_1^m \text{diag}(1, \ldots, 1, \mu^m, 1, \ldots, 1, B^m) + K_m,
\]

with \( n + 2 \) direct summands and with \( \mu^m \) in the \((m + 1)^{st}\) position. Now let \( V = V_1 V_2 \cdots V_n \), so

\[
V = \lambda \text{diag}(1, \mu, \mu^2, \ldots, \mu^n, C) + K
\]

for a unimodular complex number \( \lambda \), a bounded operator \( C \) and a compact operator \( K \). Therefore \( 0 \in W_\infty(V) \) and \( V \) is a product of \( n(n + 1)/2 \) operators that are unitarily equivalent to \( U \).

\[\square\]

3. Proof of Theorem \( \text{[8]} \)

We begin by stating a well-known result (see [12, Cor. 0.15]). Recall that \( \sigma(A) \) denotes the spectrum of an operator \( A \).

**Lemma 2.** If \( \sigma(A) \cap \sigma(B) = \emptyset \), then the operator \( \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \) is similar to \( A \oplus B \).

To prove Theorem \( \text{[8]} \) assume that \( T \) is an invertible operator which is not a scalar modulo the compacts. By a result of Brown and Pearcy [11, Theorem 2], \( T \) is similar to an operator of the form

\[
T_0 = \begin{pmatrix} 0 & A & B \\ 0 & C & D \\ 1 & E & F \end{pmatrix}
\]

acting on \( H \oplus H \oplus H \). Let \( S \) be an arbitrary invertible operator, let

\[
L_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad L_2 = \begin{pmatrix} S & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},
\]

and let \( T_j = L_j^{-1} T_0 L_j \) for \( j = 1, 2 \). Then each of \( T_1 \) and \( T_2 \) is similar to \( T \) and

\[
T_2 T_1 = \begin{bmatrix} F(S) & 0 \\ \ast & S \end{bmatrix},
\]

where

\[
F(S) = \begin{bmatrix} S^{-1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} VA & VB \\ C & D \end{bmatrix} \begin{bmatrix} D & C \\ VB & VA \end{bmatrix}.
\]

For every invertible operator \( X \), we will show that \( \sigma(\alpha X) \cap \sigma(F(\alpha X)) = \emptyset \) if \( |\alpha| \) is either large enough or small enough. To prove this, notice that \( F(\alpha X) = [\alpha^{-1} 0 \ 0 \ \ast] F(X) \), so \( \|F(\alpha X)\| \leq \|F(X)\| \) for \( |\alpha| \geq 1 \) and hence we can choose \( |\alpha| \) large enough so that \( \sigma(\alpha X) \) lies outside the disc \( \{z : |z| \leq \|F(X)\|\} \) which includes \( \sigma(F(\alpha X)) \). Similarly, for \( |\alpha| \) small enough, \( \sigma(\alpha X) \) is included in the disc \( \{z : |z| < \|F(X)^{-1}\|^{-1}\} \), while \( \sigma(F(\alpha X)) \) lies outside the same disc since \( \|F(\alpha X)^{-1}\| \leq \|F(X)^{-1}\|^{-1} \).
\[ \|F(X)^{-1}\| \text{ for } |\alpha| \leq 1. \] Applying the above to \( X = S \) and \( X = 1 \) and using Lemma 2, we conclude that there exists a scalar \( \alpha \) such that each of the operators
\[
\begin{bmatrix}
F(\alpha S) & 0 \\
0 & \alpha S
\end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix}
F(\alpha^{-1}) & 0 \\
0 & \alpha^{-1}
\end{bmatrix}
\]
is a product of two operators similar to \( T \), and so \( S \oplus F(\alpha S)F(\alpha^{-1}) \) is a product of four operators similar to \( T \).

Now take \( S \) to be \( U \oplus 1 \) where \( U \) is a bilateral shift with infinite multiplicity and 1 is the identity operator on an infinite dimensional space. From the above, there exists an invertible operator \( Q \) on \( H \) such that \( S \oplus Q \) is a product of four operators similar to \( T \). The operator \( S \oplus Q \) can be written as \( U \oplus Q' \) where both \( U \) and \( Q' \) are operators on \( \sum_{n \in \mathbb{Z}} \oplus H_n \) with \( H_n = H_0 \) for all \( n \) and
\[
U(\ldots, x_{-2}, x_{-1}, [x_0], x_1, \ldots) = (\ldots, x_{-2}, x_{-1}, x_0, x_1, \ldots),
\]
\[
Q' = \text{diag}(\ldots, 1, \Box, 1, 1, \ldots);
\]
that is,
\[
Q'(\ldots, x_{-2}, x_{-1}, [x_0], x_1, x_2, \ldots) = (\ldots, x_{-2}, x_{-1}, Qx_0, x_1, x_2, \ldots).
\]
(The box \( \Box \) is used to indicate the zero\(^{th} \) position.) Now
\[
(U \oplus Q')(Q' \oplus U) = UQ' \oplus Q'U,
\]
\[
UQ'(\ldots, x_{-2}, x_{-1}, [x_0], x_1, x_2, \ldots) = (\ldots, x_{-2}, x_{-1}, Qx_0, x_1, x_2, \ldots).
\]
Let \( J = \text{diag}(\ldots, 1, 1, \Box, Q, Q, \ldots) \). By direct computation, it follows that \( J(UQ')J^{-1} = U \). In the same way, we can show that \( Q'U \) is similar to \( U \). Therefore \( (U \oplus Q')(Q' \oplus U) \) is similar to a bilateral shift of infinite multiplicity. We have shown that a bilateral shift is a product of eight operators similar to \( T \). Since each symmetry is a product of two bilateral shifts of infinite multiplicity, the theorem follows from Radjavi’s result [11] which asserts that every invertible operator is a product of at most seven involutions.

4. Groups in the Calkin algebra

The Calkin algebra \( B(H)/\mathcal{K}(H) \) will be denoted by \( \mathfrak{A} \). The group of invertible elements and unitary elements of \( \mathfrak{A} \) will be denoted by \( GL(\mathfrak{A}) \) and \( U(\mathfrak{A}) \), respectively. In this section, we make a few remarks about semigroups generated by a conjugacy class in \( GL(\mathfrak{A}) \) and \( U(\mathfrak{A}) \). Recall that two elements \( a \) and \( b \) of a group \( G \) are said to be conjugate if \( a = g^{-1}bg \) for some \( g \in G \).

Before proceeding, we recall some facts about the Calkin algebra and index theory (see [3], Chapter 5]). The index of a Fredholm operator \( T \) is defined by \( \text{ind}(T) = \dim \ker(T) - \dim \ker(T^*) \). The index satisfies the equation \( \text{ind}(TS) = \text{ind}(T) + \text{ind}(S) \). Furthermore, it is invariant under compact perturbations. Let \( \pi: B(H) \to \mathfrak{A} \) be the canonical quotient map. Atkinson’s theorem [3, Theorem 5.17] implies that the set of Fredholm operators is the inverse image under \( \pi \) of the set \( GL(\mathfrak{A}) \) of invertible operators in \( \mathfrak{A} \). In view of this and the invariance of the index under compact perturbations, we define the index of an invertible element in \( \mathfrak{A} \) by \( \text{ind}(a) = \text{ind}(A) \) for any \( A \in \pi^{-1}(a) \). This gives a homomorphism from the group \( GL(\mathfrak{A}) \) onto the group of integers \( \mathbb{Z} \).
Two facts about operators of index 0 are needed in the sequel.

1. For a Fredholm operator $T$, $\text{ind}(T) = 0$ if and only if $T$ is a compact perturbation of an invertible operator.

2. If $\pi(T)$ is unitary and if $\text{ind}(T) = 0$, then $T$ is a compact perturbation of a unitary operator [2, Theorem 3.1].

One more fact about the Calkin algebra $\mathfrak{A}$ is that the centre of $\mathfrak{A}$ is the scalars [3]. It follows immediately that the centre of the group $GL(\mathfrak{A})$ is also the (nonzero) scalars. We can also easily establish the fact that the centre of the group $U(\mathfrak{A})$ is $\{\lambda 1: |\lambda| = 1\}$ since every element of $\mathfrak{A}$ is a linear combination of four unitary elements. (Indeed, if $a$ is self-adjoint with $\|a\| \leq 1$, then $a \pm (1-a^2)^{1/2}$ are unitaries, and hence $a$ is a convex combination of two unitaries.)

We now state two immediate consequences of Theorems [A] and [B].

**Proposition 1.** Let $a$ be an invertible (respectively, a unitary) element of $\mathfrak{A}$ of index 0. If $a$ is not a scalar, then the semigroup generated by the conjugacy class of $a$ in $GL(\mathfrak{A})$ (respectively, $U(\mathfrak{A})$) is the subgroup of all elements of index 0.

**Proof.** Since $\text{ind}(a) = 0$, there exists an invertible operator $B$ such that $\pi(B) = b$. Furthermore, if $a$ is a unitary, then the operator $B$ may be chosen to be a unitary [2, Theorem 3.1]. Since $B$ is not a scalar modulo the compacts, the results follow from Theorems [A] and [B].

For more general elements, we consider only the group generated by the conjugacy class. First, we need a lemma.

**Lemma 3.** If $a$ is an invertible element in the Calkin algebra $\mathfrak{A}$ such that $a^{-1}u^{-1}au$ is a scalar for every unitary element $u$ in $\mathfrak{A}$, then $a$ is a scalar.

**Proof.** Let $b$ be a self-adjoint element in $\mathfrak{A}$. Since $e^{itb}$ is unitary for every real number $t$, there exist scalars $\lambda_t$ such that $a^{-1}e^{-itb}ae^{itb} = \lambda_t1$ for every scalar $t$. Taking the derivative at $t = 0$, we get that $b - a^{-1}ba = \lambda 1$ for a scalar $\lambda$. Thus

$$\sigma(b) = \sigma(a^{-1}ba) = \sigma(b - \lambda 1) = \sigma(b) - \lambda.$$ 

This implies that $\lambda = 0$ and hence $ab = ba$; i.e., $a$ commutes with every self-adjoint element in $\mathfrak{A}$. It follows that $a$ commutes with every element in $\mathfrak{A}$ and so $a$ is a scalar.

**Proposition 2.** Let $\mathcal{G}$ be either the group $GL(\mathfrak{A})$ of all invertible elements or the group $U(\mathfrak{A})$ of all unitary elements in the Calkin algebra. If $a$ is an element of $\mathcal{G}$ with a nonzero index $n$, then the group generated by the conjugacy class of $a$ in $\mathcal{G}$ is $\{g \in \mathcal{G}: n \text{ divides } \text{ind}(g)\}$.

**Proof.** Let $\mathcal{N}$ be the group generated by the conjugacy class of $a$. Since $a$ is not a scalar, it follows from Lemma 3 that there exists a unitary element $u$ in $\mathfrak{A}$ such that $b := a^{-1}u^{-1}au$ is not a scalar. Now $b \in \mathcal{N}$ and $\text{ind}(b) = 0$. By Proposition 1 we have that $\mathcal{N} \supseteq \mathcal{G}_0 := \{g \in \mathcal{G}: \text{ind}(g) = 0\}$. Since $\mathcal{G}_0$ is the kernel of the homomorphism $\text{ind}: \mathcal{G} \rightarrow \mathbb{Z}$, the subgroup $\mathcal{N}$ is the inverse image under the index map of a subgroup of $\mathbb{Z}$, and the result follows.

**Corollary 4.** Let $\mathcal{G}$ be either the group $GL(\mathfrak{A})$ of all invertible elements or the group $U(\mathfrak{A})$ of all unitary elements in the Calkin algebra. Every normal subgroup of $\mathcal{G}$ is either included in the centre (i.e., the scalars) or is $\{g \in \mathcal{G}: n \text{ divides } \text{ind}(g)\}$, for some integer $n$. 
Corollary 5. Let $G$ be as above, and let $S$ be the unilateral shift and $s$ its image in the Calkin algebra. Then the group generated by the conjugacy class of $s$ in $G$ is all of $G$.

References