THE NUMERICAL RANGE OF A NILPOTENT OPERATOR
ON A HILBERT SPACE

MUBARIZ T. KARAEV

(Communicated by Joseph A. Ball)

Abstract. We prove that the numerical range $W(N)$ of an arbitrary nilpotent operator $N$ on a complex Hilbert space $H$ is a circle (open or closed) with center at 0 and radius not exceeding $\|N\| \cos \frac{\pi}{n+1}$, where $n$ is the power of nilpotency of $N$.

The purpose of this note is to describe the numerical range of a nilpotent operator on a Hilbert space. Let $H$ be a complex Hilbert space, with inner product $\langle , \rangle$, and denote by $B(H)$ the algebra of bounded linear operators on $H$. The numerical range $W(T)$ of an operator $T \in B(H)$ is defined by

$$W(T) = \{ \langle Tx, x \rangle : x \in (H)_1 \},$$

where $(H)_1 = \{ x \in H : \|x\| = 1 \}$ is the unit sphere in $H$. The numerical radius $w(T)$ of an operator $T \in B(H)$ is given by

$$w(T) = \sup \{ |\lambda| : \lambda \in W(T) \}.$$

Haagerup and Harpe [HH] proved the following sharp estimate for the numerical radius of a nilpotent operator $N \in B(H)$:

$$w(N) \leq \|N\| \cos \frac{\pi}{n+1},$$

where $n$ is the power of nilpotency of $N$. By the Toeplitz-Hausdorff theorem (see [H, GR]) the numerical range $W(N)$ is convex. Then, to prove that $W(N)$ is a disk centered at zero, it suffices to show its circularity, i.e., that each point $\lambda$ is contained in $W(N)$ with the whole circle $|z| = |\lambda|$, which is essentially contained in the paper [LT] by Li and Tsing.

The present paper presents a new proof of this known result; namely, we give a different proof of circularity of $W(T)$ using the Sz.-Nagy-Foiaș model.

The main result of the paper is

**Theorem.** Let $H$ be a complex Hilbert space and let $N \in B(H)$ be a nilpotent operator with power of nilpotency $n$. Then the numerical range $W(N)$ of the operator $N$ is a circle (open or closed) with center at 0 and radius not exceeding $\|N\| \cos \frac{\pi}{n+1}$.
As noticed already, by the Toeplitz-Hausdorff theorem and by the inequality [1] of Haagerup and Harpe, it remains to establish circularity of \( W(N) \), which will be done using the Sz.-Nagy-Foiaş model. Let us recall that the characteristic function \( \Theta_T \) of the contraction \( T \in B(H) \) is defined by

\[
\Theta_T(\lambda) = \left[ -T + \sum_{m=1}^{\infty} \lambda^m D_{T^*}(T^*)^{m-1} D_T \right](\lambda \in \mathbb{D}),
\]

where

\[
D_T = (I - T^*T)^{\frac{1}{2}}, \quad D_{T^*} = (I - TT^*)^{\frac{1}{2}},
\]

and the series is norm convergent.

Let \( E_1 \) and \( E_2 \) be Hilbert spaces, \( B(E_1, E_2) \) the space of bounded linear operators from \( E_1 \) to \( E_2 \), and \( H^\infty(B(E_1, E_2)) \) the space of bounded analytic functions on the unit disk \( \mathbb{D} = \{ z \in \mathbb{C} : |z| < 1 \} \) taking values in \( B(E_1, E_2) \). We call a function \( \Theta \in H^\infty(B(E_1, E_2)) \) inner if \( \Theta(z^*) \Theta(z) = I_{E_1} \) for almost all \( z \in \mathbb{T} = \partial \mathbb{D} \), and *-inner if \( \Theta(z^*) \Theta(z)^* = I_{E_2} \) for almost all \( z \in \mathbb{T} \). A function that is inner and *-inner simultaneously is called a two-sided inner function, where \( I_{E_i} \) is an identity operator in \( E_i, i = 1, 2 \). It is known that \( \Theta_T \in H^\infty(B(D_T, D_{T^*})) \).

Recall that a contraction \( T \in B(H) \) is said [SF] to be of class \( C_0 \) if \( T \) is completely nonunitary and there is a nonzero function \( f \) in \( H^\infty \) such that \( f(T) = 0 \). Then there is a unique (up to a constant factor of modulus one) nonconstant inner function \( u \), called the minimal function of \( T \), such that \( u(T) = 0 \). It is a well-known result of Sz.-Nagy and Foiaş [SF] that each \( T \in C_0 \) is unitary equivalent to its model operator \( M_\Theta = P_\Theta S_E |K_\Theta| \), acting on the model space \( K_\Theta = H^2(E) \otimes \Theta H^2(E) \). Here \( \Theta = \Theta_T \in H^\infty(B(E)) \) is the characteristic function of the contraction \( T \) (a two-sided inner function); \( E \) is a Hilbert space with \( \dim E = \dim (I - T^*T)H^2(E) \) is the Hardy space of \( E \)-valued functions consisting of all Taylor series \( f(z) = \sum_{n=0}^{\infty} f(n) z^n, z \in \mathbb{D} \), where \( f(n) \in E, n \geq 0 \), \( \sum_{n=0}^{\infty} \| f(n) \|^2_E < +\infty \); \( P_\Theta = I - \Theta P_+ \Theta^* \) is the orthogonal projection of \( H^2(E) \) onto \( K_\Theta \), where \( P_+ \) is the Riesz projector of \( L^2(E) \) onto \( H^2(E) \), and \( S_E f = zf \) is the shift operator on \( H^2(E) \).

If \( T \in C_0 \) and \( u \in H^\infty \) is its minimal function, then by the theorem on scalar multiples (see, e.g., [SF, N]) there exists \( \Omega \in H^\infty(B(E)) \) such that

\[
\Theta_T \Omega = \Omega \Theta_T = uI_E.
\]

In particular, if \( T \) is a nilpotent contraction, \( T^n = 0, n \geq 2 \), and \( \Theta \) is its characteristic function, then there exists \( \Omega \in H^\infty(B(E)) \) such that

\[
\Theta \Omega = \Omega \Theta = z^n I_E.
\]

Therefore, it follows from the inclusion \( \Theta \Omega H^2(E) \subset \Theta H^2(E) \) that

\[
H^2(E) \otimes \Theta^2(E) \subset H^2(E) \otimes \Theta^2(E) = H^2(E) \otimes z^n H^2(E)
\]

or

\[
K_\Theta \subset H^2(E) \otimes z^n H^2(E).
\]
Consequently, every \( x \in \mathcal{K}_\Theta \) has the form

\[
(2) \quad x(z) = \sum_{m=0}^{n-1} \hat{x}(m) z^m.
\]

Then for each \( x \in (\mathcal{K}_\Theta)_1 \) we have

\[
\langle M_\Theta x, x \rangle = \langle P_\Theta zx, x \rangle = \left\langle z \sum_{m=0}^{n-1} \hat{x}(m) z^m, \sum_{m=0}^{n-1} \hat{x}(m) z^m \right\rangle = \left\langle \sum_{m=0}^{n-1} \hat{x}(m) z^{m+1}, \sum_{m=0}^{n-1} \hat{x}(m) z^m \right\rangle = \sum_{m=0}^{n-2} \left\langle \hat{x}(m), \hat{x}(m+1) \right\rangle_E
\]

(\( E \) can be identified with a subspace of constant functions in \( H^2(E) \)). That is,

\[
(3) \quad \langle M_\Theta x, x \rangle = \sum_{m=0}^{n-2} \left\langle \hat{x}(m), \hat{x}(m+1) \right\rangle_E.
\]

Thus, the numerical range \( W(M_\Theta) \) of the operator \( M_\Theta \) is the set

\[
(4) \quad W(M_\Theta) = \left\{ \sum_{m=0}^{n-2} \left\langle \hat{x}(m), \hat{x}(m+1) \right\rangle_E : x \in (\mathcal{K}_\Theta)_1 \right\}.
\]

**Proof of the Theorem.** We need only to prove the particular case of \( \|N\| \leq 1 \), and since \( \left\| \frac{N}{\|N\|} \right\| = 1 \), the general result will follow by repeated application of the particular case. Then it is obvious that \( W(N) = W(M_\Theta) \) and \( w(N) = w(M_\Theta) \), where \( M_\Theta \) is the model operator of the contraction \( N \) and \( \Theta \) is its characteristic function. Now we prove that the corresponding set \( (4) \) is a circular set. In fact, by virtue of the equality \( (2) \) we have that

\[
y_\zeta(z) \equiv \sum_{m=0}^{n-1} \zeta^m \hat{x}(m) z^m \in \mathcal{K}_{z^n I_E}
\]

for each \( x = \sum_{m=0}^{n-1} \hat{x}(m) z^m \in \mathcal{K}_\Theta \) and \( \zeta \in \mathbb{T} \), and therefore, for all \( h \in H^2(E) \),

\[
0 = \langle y_\zeta, z^n h \rangle = \langle y_\zeta, \Omega \Theta h \rangle = \langle \Omega^* y_\zeta, \Theta h \rangle = \langle P_+ \Omega^* y_\zeta, \Theta h \rangle,
\]

which implies that \( \Omega^* y_\zeta \in L^2(E) \otimes \Theta H^2(E) \) and \( P_+ \Omega^* y_\zeta \in \mathcal{K}_\Theta \). Since \( L^2(E) \otimes \Theta H^2(E) = \mathcal{K}_\Theta \otimes H^2_-(E) \), we have that \( \Omega^* y_\zeta = P_+ \Omega^* y_\zeta + h_- \), where \( h_- \in H^2_-(E) \).
Let $x \in W(M_T)$ be realized in the element $x = \sum_{m=0}^{n-1} \hat{x}(m) z^m \in (K_1^T)^*$, i.e., $\lambda = \langle M_T x, x \rangle$. Then, by virtue of (3), for any $t \in [0, 2\pi]$ we have

$$e^{it} \lambda = e^{it} \langle M_T x, x \rangle = e^{it} \sum_{m=0}^{n-2} \left\langle \hat{x}(m), \hat{x}(m+1) \right\rangle_E$$

$$= \sum_{m=0}^{n-2} \left\langle e^{-imt} \hat{\lambda}, e^{-i(m+1)t} \hat{\lambda} \right\rangle_E$$

$$= \sum_{m=0}^{n-2} \left\langle \hat{y}_\lambda(m), \hat{y}_\lambda(m+1) \right\rangle_E = (zy_\lambda, y_\lambda),$$

where

$$y_\lambda(z) = \sum_{m=0}^{n-1} \hat{y}_\lambda(m) z^m \in (K_{z^n}^T) \cap E,$$

$$\hat{y}_\lambda(m) = \sum_{m=0}^{n-1} \hat{x}(m) e^{-imt} \hat{x}(m) (m \geq 0).$$

On the other hand,

$$\langle M_T P_+ \Omega^* y_\lambda, P_+ \Omega^* y_\lambda \rangle = \langle P_+ \Omega^* y_\lambda, \Omega^* P_+ \Omega^* y_\lambda \rangle = \langle zP_+ \Omega^* y_\lambda, P_+ \Omega^* y_\lambda \rangle$$

$$= \langle zP_+ \Omega^* y_\lambda, \Omega^* y_\lambda \rangle = \langle z(\Omega^* y_\lambda - \langle h_-, \Omega^* y_\lambda \rangle)$$

$$= \langle z\Omega^* y_\lambda, \Omega^* y_\lambda \rangle - \langle h_-, \Omega^* y_\lambda \rangle$$

$$= \langle y_\lambda, \Omega^* y_\lambda \rangle - \langle h_-, \Omega^* y_\lambda \rangle$$

$$= e^{it} \langle M_T x, x \rangle - \langle h_-, \Omega^* y_\lambda \rangle = e^{it} \lambda - \langle h_-, \Omega^* y_\lambda \rangle.$$

Thus

$$e^{it} \lambda = \left\| P_+ \Omega^* y_\lambda \right\|^2 \left\langle M_T \frac{P_+ \Omega^* y_\lambda}{\left\| P_+ \Omega^* y_\lambda \right\|}, \frac{P_+ \Omega^* y_\lambda}{\left\| P_+ \Omega^* y_\lambda \right\|} \right\rangle + \langle h_-, \Omega^* y_\lambda \rangle,$$

or

$$e^{it} \lambda = \left( 1 - \left\| h_- \right\|^2 \right) \left\langle M_T \frac{P_+ \Omega^* y_\lambda}{\left\| P_+ \Omega^* y_\lambda \right\|}, \frac{P_+ \Omega^* y_\lambda}{\left\| P_+ \Omega^* y_\lambda \right\|} \right\rangle + \langle h_-, \Omega^* y_\lambda \rangle. \quad (5)$$

We shall prove that there exists $x_0 \in (K_1^T)^*$ such that

$$\langle h_- \Omega^* y_\lambda \rangle = \left\| h_- \right\|^2 \langle M_T x_0, x_0 \rangle,$$

and so formulas (3) and (6) will give the desired inclusion $e^{it} \lambda \in W(M_T)$, since $\left\| h_- \right\|^2 < 1$ and $W(M_T)$ is convex.

First recall that if $F_1, F_2$ are two subspaces of $H$, the angle between $F_1$ and $F_2$ is the number $\alpha = \alpha_{F_1, F_2}, 0 \leq \alpha \leq \frac{\pi}{2}$, such that

$$\cos \alpha = \sup_{x_1 \in F_1, x_2 \in F_2} \frac{\left\langle x_1, x_2 \right\rangle}{\left\| x_1 \right\| \left\| x_2 \right\|}.$$

Since ker $M_T \neq \{0\}$, then $0 \in W\left( \left\| h_- \right\|^2 M_T \right)$. From the convexity of numerical range it follows that $W\left( \left\| h_- \right\|^2 M_T \right)$ contains the line segment $[0, \left\| h_- \right\|^2 \lambda]$, and moreover, $\Sigma_r \triangleq \{ z : |z| \leq r \} \subset W\left( \left\| h_- \right\|^2 M_T \right)$ for some $r, 0 < r < \left\| h_- \right\|^2 |\lambda|$.

Let $L_1 \subset L^2(E)$ and $L_2 \subset L^2(E)$ be the subspaces such that $zh_- \in L_1$ and $\Omega^* y_\lambda \in L_2$. Recall that two contractive functions $\Theta_1 \in H^\infty(B(E_1, E_2)), \Theta_2 \in$
$H^\infty (B (E'_1, E'_2))$ coincide (see [4], Chapter 5) if there exist unitary operators $	au_1 : E_1 \to E'_1$ and $	au_2 : E_2 \to E'_2$ such that $\Theta_2 (z) = \tau_2 \Theta_1 (z) \tau_1^{-1}, |z| < 1$. It is well known that two completely nonunitary contractions are unitarily equivalent if and only if their characteristic functions coincide. Therefore, we can assume the function $\Omega$ and the subspaces $L_1, L_2$ are such that $\|h_\cdot\| \cos \alpha_{L_1, L_2} \leq r$. Then we have

$$\|h_\cdot\| \frac{|\langle z h_\cdot, \Omega^* y_\zeta \rangle|}{\|z h_\cdot\| \|\Omega^* y_\zeta\|} \leq \|h_\cdot\| \sup_{x_1 \in L_1, x_2 \in L_2} \frac{|\langle x_1, x_2 \rangle|}{\|x_1\| \|x_2\|}$$

$$= \|h_\cdot\| \cos \alpha_{L_1, L_2} \leq r,$$

which shows that $\langle z h_\cdot, \Omega^* y_\zeta \rangle \in D_r$. Since $D_r \subset W (\|h_\cdot\|^2 M_0)$, there exists $x_0 \in (K_\Omega)_1$ such that

$$\langle z h_\cdot, \Omega^* y_\zeta \rangle = \langle \|h_\cdot\|^2 M_0 x_0, x_0 \rangle = \|h_\cdot\|^2 \langle M_0 x_0, x_0 \rangle.$$

Consequently, $e^{it\lambda} \in W (M_0)$ for each $t \in [0, 2\pi]$, which means that $W (M_0)$ is a circular set. On the other hand, since the numerical range is convex, $W (M_0)$ is a circle (open or closed) with center at zero, which completes the proof.

**Corollary.** If $N \in B (H)$ is a compact nilpotent operator, then $W (N)$ is a closed disk centered at 0 and with radius not greater than $\|N\| \cos \frac{\pi}{n+1}$.

**Proof.** It follows from the result of Lancaster [5] that if $K \in B (H)$ is compact, then $W (K)$ is closed if and only if $0 \in W (K)$, which completes the proof.

**Example 1.** Let $(V_0 f) (x) = \int_{-x}^x f (t) \, dt$ be a skew-symmetric Volterra operator acting on the space $L^2 (-1, 1)$. Then its numerical range $W (V_0)$ is a closed circle with center at 0 and radius $\frac{\pi}{2}$, i.e., $W (V_0) = D_{\frac{\pi}{2}}$.

**Proof.** Observe that the set of values of the operator $V_0$ is contained in the set of all odd functions from the space $L^2 (-1, 1)$ and for the odd function $f$, $V_0 f = 0$. This means that $V_0^2 = 0$, i.e., $V_0$ is a nilpotent operator with the power of nilpotency 2. Then by our theorem, $W (V_0)$ is a circle with center at zero. Since $V_0$ is a compact operator and $0 \in W (V_0)$, by the corollary, $W (V_0)$ is a closed set. So, $W (V_0)$ is a closed circle. Now we calculate its radius. By inequality (1), $w (V_0) \leq \frac{\|V_0\|}{2}$. On the other hand, $w (V_0) \geq \frac{\|V_0\|}{2}$, and hence $w (V_0) = \frac{\|V_0\|}{2}$. Since $\|V_0\| = \frac{4}{\pi}$ (see [6]), we have that $w (V_0) = \frac{2}{\pi}$. This completes the proof. 

The following example is easy to verify.

**Example 2.** Consider the diagonal operator $D = \begin{bmatrix} 0 & D \\ 0 & 0 \end{bmatrix}$ and $N = \begin{bmatrix} 0 & D \\ 0 & 0 \end{bmatrix}$ acting in $l^2$. Then $N$ is not compact, $N^2 = 0$, and $W (N)$ is the open disk centered at 0 with radius $\frac{\pi}{2}$, i.e., $W (N) = D_{\frac{\pi}{2}}$. 


**Remark.** Recall that the $k$-numerical range $W_k(A)$ of the operator $A \in B(H)$ is defined by

$$W_k(A) = \{ \lambda \in \mathbb{C} : \lambda = \sum_{i=1}^{k} \langle Ax_i, x_i \rangle, \}$$

where $\{x_1, \ldots, x_k\}$ are $k$ orthonormal vectors in $H$.

When $k = 1$, this reduces to the classical numerical range $W(A)$. The most fundamental property of the $k$-numerical range is its convexity (see [H]). Using this property and arguments similar to those in the proof of the theorem, we can also obtain the analogous characterization for the $k$-numerical range $W_k(N)$ of a nilpotent operator $N \in B(H)$, which is also contained among other general results of the paper [LT].

I am grateful to the referee for his constructive suggestions and for pointing out the reference [LT].

**REFERENCES**


Institute For Mathematics And Mechanics, Azerbaijani National Academy of Sciences, F. Agaev, 9, 370141 Baku, Azerbaijan

Current address: Department of Mathematics, Faculty of Arts and Sciences, Suleyman Demirel University, 32260 Isparta, Turkey

E-mail address: garayev@fef.sdu.edu.tr