Časopis pro pěstování matematiky, roč. 103 (1978), Praha

UNIQUENESS OF THE OPERATOR ATTAINING $C(H_n, r, n)$

ZDENĚK DOSTÁL, Ostrava

(Received May 24, 1976)

Introduction. Let r be a fixed real number, $0 < r < 1$, n a fixed natural number. Let $L(H_n)$ denote the algebra of all linear operators on an n-dimensional Hilbert space H_n and let the operator norm and the spectral radius of $A \in L(H_n)$ be denoted by $|A|$ and $|A|_\sigma$, respectively.

In connection with the critical exponent, V. PTÁK has introduced in [1] the quantity

$$C(H_n, r, m) = \sup \{|A^m| : A \in L(H_n),\ |A|_\sigma \leqq r,\ |A| \leqq 1\}$$

and found a certain operator $A \in L(H_n)$ such that

$$(1) \qquad C(H_n, r, n) = |A^n|\,, \quad |A|_\sigma \leqq r\,, \quad |A| \leqq 1\,.$$

The point of this note is to show that the operator A is unique in the following sense: if $B \in L(H_n)$ is any operator which satisfies (1) then there exists a unitary operator $U \in L(H_n)$ and a complex unit ε such that

$$\varepsilon A = U^* B U\,.$$

2. Notation and preliminaries. Let M_n denote the algebra of all $n \times n$ complex valued matrices.

The adjoint and the spectrum of an operator A will be denoted by A^* and $\sigma(A)$, respectively.

An operator $A \in L(H_n)$ is said to be extremal if $|A| \leqq 1$, $|A|_\sigma \leqq r$ and $|A^n| = C(H_n, r, n)$.

For a given set $W = \{w_1, \ldots, w_n\}$ of vectors $w_i \in H_n$, denote by $G(W)$ the Gramm matrix of W. If $z \in H_n$ and $A \in L(H_n)$, we shall abbreviate $G(z, Az, \ldots, A^{n-1}z)$ by $G(A, z)$.

We shall denote, for $1 \leqq i \leqq n$, by E_i the polynomial

$$E_i(x_1, \ldots, x_n) = \sum_{\substack{e_j \in \{0,1\} \\ e_1 + \ldots + e_n = i}} x_1^{e_1} x_2^{e_2} \ldots x_n^{e_n}\,,$$

Let $\varrho_1, \ldots, \varrho_n$ be given complex numbers. For $i = 1, 2, \ldots, n$, put $\alpha_i = (-1)^{n-i} E_{n-i+1}(\varrho_1, \ldots, \varrho_n)$ so that the roots of the equation

$$x^n = \alpha_1 + \alpha_2 x + \ldots + \alpha_n x^{n-1}$$

are exactly $\varrho_1, \ldots, \varrho_n$. Consider the recursive relation

$$x_{i+n} = \alpha_1 x_i + \ldots + \alpha_n x_{i+n-1} . \tag{2}$$

For each i, $1 \leqq i \leqq n$, we denote by $w_i(\varrho_1, \ldots, \varrho_n)$ the solution $(w_{i0}, w_{i1}, w_{i2}, \ldots)$ of this relation with the initial conditions

$$w_{ik}(\varrho_1, \ldots, \varrho_n) = \delta_{i,k+1} , \quad 0 \leqq k \leqq n - 1 .$$

The result of V. KNICHAL ([1], Lemma 7) reads:

2.1. *For each $i = 1, 2, \ldots, n$ and each $k \geqq n$,*

$$w_{ik}(\varrho_1, \ldots, \varrho_n) = \varepsilon_i Q_{ik}(\varrho_1, \ldots, \varrho_n) ,$$

where $\varepsilon_i = (-1)^{n-i}$ and

$$Q_{ik}(\varrho_1, \ldots, \varrho_n) = \sum_{\substack{e_j \geqq 0 \\ e_1 + \ldots + e_n = k - i + 1}} c_{ik}(e_1, \ldots, e_n)\, \varrho_1^{e_1}, \ldots, \varrho_n^{e_n} ,$$

where all $c_{ik}(e_1, \ldots, e_n) \geqq 0$.

The point of the lemma is that, for $k \geqq n$ and i fixed, all coefficients of w_{ik} are of the same sign.

Following [1], we denote by $P(\varrho_1, \ldots, \varrho_n)$ the linear space consisting of all solutions of the recursive relation (2); it is spanned by the vectors $w_1(\varrho_1, \ldots, \varrho_n), \ldots, w_n(\varrho_1, \ldots \ldots, \varrho_n)$.

Now suppose that all $|\varrho_i| < r$. It is proved in [1] that, in this case, $P(\varrho_1, \ldots, \varrho_n)$ is a subspace of the Hilbert space l^2 of all sequences $(a_0, a_1, a_2, \ldots)$ of the complex numbers such that $\sum_{i=0}^{\infty} |a_i|^2 < \infty$.

Let S denote the shift operator on l^2 which sends $(a_0, a_1, a_2, \ldots)$ to $(a_1, a_2, a_3, \ldots)$. Its restriction on $P(\varrho_1, \ldots, \varrho_n)$ is denoted by $S \mid P(\varrho_1, \ldots, \varrho_n)$.

The solution $(a_0, a_1, a_2, \ldots)$ of (2) with the initial conditions $a_0 = 1$, $a_1 = \varrho_i, \ldots \ldots, a_{n-1} = \varrho_i^{n-1}$ is the eigenvector corresponding to ϱ_i. On the other hand,

$$(S^n - \alpha_n S^{n-1} - \ldots - \alpha_1) \mid P(\varrho_1, \ldots, \varrho_n) = 0$$

so that the minimal polynomial of $S \mid P(\varrho_1, \ldots, \varrho_n)$ is a divisor of $(x - \varrho_1) \ldots \ldots (x - \varrho_n)$. We have thus

$$\sigma(S \mid P(\varrho_1, \ldots, \varrho_n)) = \{\varrho_1, \ldots, \varrho_n\} . \tag{3}$$

3. Shifts. V. Pták has discovered extremal properties of restrictions of the shift S. He has proved:

3.1. Theorem. (Pták). *Let $\varrho_1, \ldots, \varrho_n$ be complex numbers, $|\varrho_i| \leqq r$ for $i = 1, \ldots, n$; $A \in L(H_n)$, $|A| \leqq 1$ and $(A - \varrho_1)(A - \varrho_2) \ldots (A - \varrho_n) = 0$.*

Then

$$|A^n| \leqq |S^n| P(\varrho_1, \ldots, \varrho_n)| \tag{4}$$

([1], Theorem 6).

Moreover,

$$C(H_n, r, n) = |S^n | P(r, \ldots, r)| \tag{5}$$

(ibid, Theorem 8).

The proof of (5) consists in showing that

$$|S^n | P(\varrho_1, \ldots, \varrho_n)| \leqq |S^n | P(r, \ldots, r)| . \tag{6}$$

An inspection of the proof of (5) suggests a supplement to the inequality (6).

3.2. *Let $\varrho_1, \ldots, \varrho_n$ be complex numbers, $|\varrho_i| \leqq r$ for $i = 1, \ldots, n$. Then the relation*

$$|S^n | P(\varrho_1, \ldots, \varrho_n)| = |S^n | P(r, \ldots, r)|$$

holds if and only if $\varrho_1 = \ldots = \varrho_n$ and $|\varrho_1| = r$.

We shall follow [1] in the proof.

Let Q_i, w_i and E_i be those of Section 2. With the aid of the recurrent relations (2), it is easy to verify directly that

$$Q_{in} = E_{n-i+1} \quad \text{and} \quad Q_{1,n+1} = E_1 \,.\, E_n \,.$$

Now suppose all $|\varrho_i| \leqq r$ and let there be i such that $\varrho_1 \neq \varrho_i$ or $|\varrho_i| < r$. It follows immediately that

$$|Q_{1,n+1}(\varrho_1, \ldots, \varrho_n)| < Q_{1,n+1}(r, \ldots, r) \tag{7}$$

and

$$|Q_{i,n}(\varrho_1, \ldots, \varrho_n)| < Q_{i,n}(r, \ldots, r) \,, \quad i = 2, \ldots, n \,. \tag{8}$$

All coefficients of the forms Q_{ik} being nonnegative, we have

$$|Q_{ik}(\varrho_1, \ldots, \varrho_n)| \leqq Q_{ik}(r, \ldots, r) \,, \quad i = 1, \ldots, n \,. \tag{9}$$

We intend to show that

$$|S^n | P(\varrho_1, \ldots, \varrho_n)| < |S^n | P(r, \ldots, r)| \,.$$

To prove this, we associate with each $x \in P(\varrho_1, \ldots, \varrho_n)$, $x \neq 0$, a vector $y \in P(r, \ldots, r)$ such that

$$|S^n x|\,|x|^{-1} < |S^n y|\,|y|^{-1}\,.$$

Put $y = \sum_{i=1}^{n} |x_{i-1}|\,(-1)^{n-i}\,w_i(r, \ldots, r)$. It follows that, for $0 \leqq k \leqq n-1$, we have $|x_k| = |y_k|$. If $k \geqq n$, then

$$(10) \qquad |x_k| = \left|\sum_{i=1}^{n} x_{i-1}\,w_{ik}(\varrho_1, \ldots, \varrho_n)\right| \leqq \sum_{i=1}^{n} |x_{i-1}|\,|Q_{ik}(\varrho_1, \ldots, \varrho_n)| \leqq$$

$$\leqq \sum_{i=1}^{n} |x_{i-1}|\,Q_{ik}(r, \ldots, r) = \sum_{i=1}^{n} y_{i-1}(-1)^{n-i}\,Q_{ik}(r, \ldots, r) = y_k\,.$$

If $x_0 \neq 0$, then we can apply the inequality (7) together with (9) to get $|x_{n+1}| < y_{n+1}$, otherwise by (8) $|x_n| < y_n$. We have thus $|x_k| = |y_k|$ for $k = 0, 1, \ldots, n-1$; $|x_k| \leqq y_k$ for $k \geqq n$, $|x_n| < y_n$ or $|x_{n+1}| < y_{n+1}$ and this implies the desired inequality.

On the other hand, if $\varrho = e^{it}r$, t real, then by (6) and (4)

$$|S^n \mid P(\varrho, \ldots, \varrho)| \leq |S^n \mid P(r, \ldots, r)| = |(e^{it}S)^n \mid P(r, \ldots, r)| \leq |S^n \mid P(\varrho, \ldots, \varrho)|\,,$$

which completes the proof.

We shall need a little more information about $S \mid H(\varrho, \ldots, \varrho)$. Let $|\varrho| < 1$, and abbreviate $S \mid P(\varrho, \ldots, \varrho)$ by S_ϱ, $w_n(\varrho, \ldots, \varrho)$ by w. Clearly $|w| = |S_\varrho w| = \ldots$ $\ldots = |S_\varrho^{n-1} w|$. All the vectors $w, S_\varrho w, \ldots, S_\varrho^{n-2} w$ being linearly independent eigenvectors of $S_\varrho^* S_\varrho \neq I$ corresponding to the eigenvalue 1, we have

$$(11) \qquad \operatorname{rank}(I - S_\varrho^* S_\varrho) = 1\,.$$

We intend to show that $|S_\varrho^n z|$ attains its maximum on the unit sphere for a unique vector. To prove it, assume $u, v \in P(\varrho, \ldots, \varrho)$ linearly independent, $|u| = |v| = 1$, $|S_\varrho^n u| = |S_\varrho^n v| = |S_\varrho^n|$, i.e. $|S_\varrho^n|^2 = |S_\varrho^{*n} S_\varrho^n| = (S_\varrho^{*n} S_\varrho^n u, u) = (S_\varrho^{*n} S_\varrho^n v, v)$. It follows that both u and v are eigenvectors of $S_\varrho^{*n} S_\varrho^n$ corresponding to the eigenvalue $|S_\varrho^n|^2$ and, consequently, $|S_\varrho^n|^2\,|z|^2 = (S_\varrho^{*n} S_\varrho^n z, z) = |S_\varrho^n z|^2$ for each $z \in \operatorname{Span}(u, v)$. Since $\dim \operatorname{Ker}(I - S_\varrho^* S_\varrho) = n - 1$ and S_ϱ is regular there exists a nonzero w, $w \in S^n(\operatorname{Span}(u, v)) \cap \operatorname{Ker}(I - S_\varrho^* S_\varrho)$. Setting $z = |S_\varrho^{-n} w|^{-1}\,S_\varrho^{-n} w$ we have

$$(12) \qquad (I - S_\varrho^* S_\varrho)\,S_\varrho^n z = 0\,, \quad |S_\varrho^n z| = |S_\varrho^n| = C(H_n, r, n)\,.$$

Hence we can write

$$(13) \qquad |S_\varrho^n z|^2 - |S_\varrho^{n+1} z|^2 = ((I - S_\varrho^* S_\varrho)\,S_\varrho^n z, S_\varrho^n z) = 0\,.$$

Now return to the proof of 3.2 and set $y = \sum_{i=1}^{n} z_{i-1}(-1)^{n-i}\,.\,w_i(r, \ldots, r)$. We have again $|z_i| = |y_i|$ for $i = 0, 1, \ldots, n-1$ and $|z_i| \leqq y_i$ for $i = n, n+1, \ldots$. Applying (12) we get even $|z_i| = y_i$ for $i \geqq n$. Since by (13) $|S_\varrho^n z| = |S_\varrho^{n+1} z|$, we have $z_n = 0$.

At the same time

$$|z_n| = y_n = \sum_{i=1}^{n} |z_{i-1}| \, Q_{in}(r, \ldots, r) = \sum_{i=1}^{n} |z_{i-1}| \, E_{n-i+1}(r, \ldots, r) > 0 \,,$$

which is impossible. We have proved the following result:

3.3. *Let $|\varrho| < 1$, $u, v \in P(\varrho, \ldots, \varrho)$, $|u| = |v| = 1$ and $|S^n u| = |S^n v| = C(H_n, r, n)$. Then $u = e^{it} v$.*

4. Spectrum of extremal operators. Now it is easy to describe the spectrum of extremal operators.

4.1. *If $A \in L(H_n)$ is extremal, then $\sigma(A) = \{\varrho\}$, $|\varrho| = r$.*

Proof. Suppose $\varrho_1, \ldots, \varrho_n$ are the roots of the characteristic polynomical of an extremal operator $A \in L(H_n)$. If they were not all equal or some $|\varrho_i| < r$, then, since $(A - \varrho_1) \ldots (A - \varrho_n) = 0$, by 3.1 a 3.2

$$|A^n| \leq |S^n \mid P(\varrho_1, \ldots, \varrho_n)| < |S^n \mid P(r, \ldots, r)| = C(H_n, r, n) \,.$$

We shall need two easy consequences of 4.1.

4.2. *If $A \in L(H_n)$ is extremal, $z \in H_n$, $|z| = 1$ and $|A^n z| = A^n$, then the vectors z, $Az, \ldots, A^{n-1} z$ are linearly independent.*

Really, otherwise we could define an extremal operator B which has 0 in its spectrum by setting $Bx = Ax$ for x from the linear span of the vectors $z, Az, \ldots, A^{n-1} z$ and $By = 0$ on the orthogonal complement.

It follows that no extremal operator can be a root of the polynomial of a degree less than the dimension of the space. Together with 4.1, this yields

4.3. *If $A \in L(H_n)$ is extremal then its minimal polynomial is $(x - \varrho)^n$, where $|\varrho| = r$.*

5. We give a brief account of Pták's method of linearization that we need here ([1], pp. 250–253). In the sequel, let $z \in H_n$ be a fixed unit vector, $\varrho = e^{it} r$ a fixed real number and let T be the companion matrix of $(x - \varrho)^n$, that is

$$T = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \ldots & \alpha_n \end{bmatrix},$$

where α_i are defined by

$$(x - \varrho)^n = x^n - \alpha_n x^{n-1} - \ldots - \alpha_1 \,.$$

If $A \in L(H_n)$ satisfies $(A - \varrho)^n = 0$, then it is easy to verify directly that for each $z \in H_n$

$$G(A, Az) = TG(A, z)\, T^* . \tag{14}$$

We denote by $\mathscr{A}$ the class of all operators $A \in L(H_n)$ such that $|A| \leqq 1$ and $(A - \varrho)^n = 0$, by $\mathscr{Z}$ the class of all symmetric matrices $Z \in M_n$ satisfying $TZT^* \leqq Z$ and $z_{11} = 1$. The mapping

$$g_z : \mathscr{A} \ni A \mapsto G(A, z) \in \mathscr{Z}$$

is epimorphic.

The crucial point is that there is a linear isomorphism between the cone $\mathscr{T}$ of all symmetric matrices $Z \in M_n$, $TZT^* \leqq Z$, and the cone $\mathscr{P}$ of all symmetric positive semidefinite matrices. It is defined by

$$p : \mathscr{T} \ni Z \mapsto Z - TZT^* \in \mathscr{P} .$$

Let us define a linear functional

$$f : M_n \ni Z \mapsto q(T^n Z T^{*n}) ,$$

where $q(Z)$ denotes the (1,1) entry of Z, and let $\mathscr{Q} = p(\mathscr{Z})$. If $A \in \mathscr{A}$, we may write

$$fp^{-1}(p\, g_z(A)) = f(g_z(A)) = |A^n z|^2 ,$$

so that max $|A^n z|^2$ for $A \in \mathscr{A}$ equals the maximum of fp^{-1} on the set $\mathscr{Q}$. The last set being compact and convex, the maximum of fp^{-1} will be attained at an extreme point of $\mathscr{Q}$. Since the extreme rays of $\mathscr{P}$ are generated by matrices of the rank 1, the rank of the extreme matrices of $\mathscr{Q}$ is equal to 1.

Put $\mathscr{E} = \{P \in \mathscr{Q} : fp^{-1}(P) = C(H_n, r, n)^2\}$. First we show what do the operators from $\mathscr{A}$, which are sent by pg_z to the extremal point of $\mathscr{E}$, look like.

5.1. *Let $A \in L(H_n)$ be extremal. If the rank of the matrix*

$$G(A, z) - G(A, Az)$$

is equal to 1 and $|A^n z| = C(H_n, r, n)$, then there is a complex number ϱ, $|\varrho| = r$ and a unitary mapping

$$u : H_n \to P(\varrho, \ldots, \varrho)$$

such that

$$A = u^* S u .$$

Proof. Suppose A satisfies the assumptions of the theorem and put $D = (I - A^*A)^{1/2}$. We have seen already that $\sigma(A) = \{\varrho\}$, $|\varrho| = r$. Obviously,

$$G(A, z) - G(A, Az) = G(Dz, DAz, \ldots, DA^{n-1}z) .$$

By 4.2 the vectors $z, Az, \ldots, A^{n-1}z$ form a basis of the space H_n. The rank of $G(Dz, \ldots, DA^{n-1}z)$ being equal to 1, the same holds for D, too.

We denote by e the only unit eigenvector of D with the eigenvalue different from zero and define a linear mapping

$$u : H_n \ni w \mapsto ((Dw, e), \ (DAw, e), \ldots) \in l^2 .$$

Clearly u maps H_n into $P(\varrho, \ldots, \varrho)$. Since $A^n \to 0$ and $Dw = (Dw, e)\, e$, we have

$$|u(w)|^2 = \sum_{i=0}^{\infty} |(DA^i w, e)|^2 = \sum_{i=0}^{\infty} |DA^i w|^2 = \sum_{i=0}^{\infty} (|A^i w|^2 - |A^{i+1} w|^2) = |w|^2$$

so that u is an isometry. The spaces H_n and $P(\varrho, \ldots, \varrho)$ having the same dimension n, the range of u is $P(\varrho, \ldots, \varrho)$. Moreover, the shift S satisfies

$$uA = Su ,$$

which completes the proof.

The next step consists in showing that $\mathscr{E}$ is a singleton. To prove it, assume P, Q are extreme points of $\mathscr{E}$ and let $A, B \in \mathscr{A}$ be such operators that $p\, g(A) = P, p\, g(B) = Q, |A^n z| = |B^n z| = C(H_n, r, n)$.

By 5.1 there are isometries $u, v : H_n \to P(\varrho, \ldots, \varrho)$,

$$A = u^* S u , \quad B = v^* S v .$$

It immediately follows that

$$|S^n uz| = |S^n vz| = |A^n z| = C(H_n, r, n) ,$$

by 3.3 we get $uz = e^{it} vz$ and clearly $z = e^{-it} v^* uz$. The desired relation

$$P = p\, g(A) = p\, g(B) = Q$$

is now an easy consequence of $B = v^* u A u^* v$.

Now, if A is any extremal operator, then there is $z \in H_n$ such that $|z| = 1$ and $|A^n z| = C(H_n, r, n)$. Clearly $p\, g_z(A) \in \mathscr{E}$. Since the only matrix belonging to $\mathscr{E}$ is of rank 1, the rank of

$$p\, g_z(A) = G(A, z) - G(A, Az)$$

is equal to 1 and A satisfies the assumptions of 5.1.

We can summarize our results in the promised theorem.

5.2. Theorem. *Let $A \in L(H_n)$, $|A| \leq 1$, $0 < r < 1$, $|A|_\sigma \leq r$ and $|A^n| = C(H_n, r, n)$.*

Then $\sigma(A)$ consists of an only point ϱ, $|\varrho| = r$ and A is unitary similar to the restriction of the shift operator S on the space of all sequences $(x_0, x_1, x_2, \ldots)$ which satisfy

$$\sum_{i=0}^{n} \binom{n}{i} (-\varrho)^i x_{k+n-i} = 0 .$$

The problem of uniqueness of extremal operators was raised by V. Pták.