Commentationes Mathematicae Universitatis Carolinae

N. J. Young

Norms of certain rational functions of a matrix and Schur's criterion for polynomials

Commentationes Mathematicae Universitatis Carolinae, Vol. 19 (1978), No. 4, 673--688

Persistent URL: http://dml.cz/dmlcz/105883

COMMENTATIONES MATHEMATICAE UNIVERSITATIS CAROLINAE

19,4 (1978)

NORMS OF CERTAIN RATIONAL FUNCTIONS OF A MATRIX AND SCHUR'S CRITERION FOR POLYNOMIALS

N.J. YOUNG, Praha

Abstract: A theorem of Schur asserts that all the roots of a polynomial p are of modulus less than one if and only if a certain Hermitian form derived from p is positive definite. It is shown here that there is a wide class of Hermitian forms with the same property. The proof depends on the following fact: if g is a rational function of the form $g(z) = p(z)/z^n p(1/z)^-$ and A is an $n \times n$ matrix satisfying rank $(I - A^* A) = 1$, $\|A\| \leq 1$ and having spectral radius less than one, then $\|g(A)\| < 1$ or $\|g(A)\| > 1$ depending on whether or not all roots of p lie in the open unit disc ($\| \cdot \|$ denotes the operator norm on n-dimensional Hilbert space).

Key words: Norm, spectral radius, polynomial, root, Hermitian form.

AMS: Primary 15A42, 12D10
Secondary 46C10, 34D99

In 1917 I. Schur wrote a paper [5] on a subject which at that time interested many mathematicians, namely, the relation between the H^∞-norm of a function and its Taylor coefficients. Schur's paper is still cited for two reasons: firstly, it is the source of the formula for the determinant of a partitioned matrix which is known as Schur's formula, and secondly, it contains his well-known criterion for a polynomial to have all its roots in the open unit disc. These were, however, merely a lemma and a corollary of the main results, and the paper must be judged to have led to a dead end. Apart from a slight further development by Nevanlinna [1], no further use seems to have been made of the algebraic techniques developed by Schur. And indeed, many facts obtained in the paper by laborious algebraic computation can be seen much more easily and naturally by means of the geometric and analytic approaches now familiar to us, but which were in their infancy in 1917. For example, after fifteen pages of mighty wielding of determinants Schur arrives at a result (Theorem X) expressible in modern terminology as follows: if g is a formal power series then the norm of the operation of multiplication by g, thought of as an operator on the Hardy space H^2 of the open unit disc, is equal to the H^∞-norm of g. Nowadays this fact is almost obvious: it can be proved in a few lines with the aid of the Poisson kernel. But all this is not to say that Schur has been superseded entirely, for his methods also yield at the same time facts which are not at all so obvious to modern eyes. Such are, roughly speaking, finite dimensional versions of results of the above type holding for special classes of functions g. Here is an illustration. Let g be a given by a formula

$$g(z) = \frac{p(z)}{z^n p(1/\bar{z})^-}$$

where p is a polynomial of degree n, and let S_n be the shift operator on n-dimensional Hilbert space:

$$S_n(x_1,\dots,x_n) = (x_2,\dots,x_n,0).$$

Then g is bounded in the open unit disc if and only if $g(S_n)$ is a contraction. The criterion for polynomials mentioned above is a corollary of this fact.

Generally speaking questions about operators are easier in finite than in infinite dimensions, but here we have to deal with questions which gain their interest from finite dimensionality, and are trivial in the infinite-dimensional case. In view of the fact that our understanding of Schur's main results gains so much in simplicity, clarity and generality from a functional analytic approach, it is natural to look for a similarly advantageous treatment of problems of the latter sort. This article makes a contribution to this programme.

Problems of such a nature have in fact been handled geometrically, but only relatively recently. One of the earliest examples is the paper of V. Pták [2], where it is proved that if A is a contraction on n-dimensional Hilbert space having spectral radius less than one then A^n has norm strictly less than one. Several other proofs of this fact have been published since: references to them can be found in [3]. We begin with a generalization of Pták's theorem. The symbol $\| \cdot \|$ will denote both the norm of an element of a Hilbert space H and the operator norm of an operator on H, while $|A|_\sigma$ will denote the spectral radius of an operator A on H.

<u>Theorem 1</u> Let $A_1, A_2, \ldots, A_n$ <u>be commuting linear operators on an</u> n-<u>dimensional Hilbert space</u> H, <u>and let</u> $|A_i|_\sigma < 1$ <u>and</u> $\|A_i\| \leq$ <u>for each</u> i. <u>Then</u> $\|A_1 A_2 \ldots A_n\| < 1$.

This result has already been proved and used [6], but the proof is very short and is worth repeating since it should make the sequel easier to follow.

Let us first recall the fact that if an operator A on H satisfies $\|A\| \leq 1$, then $\{x: \|Ax\| = \|x\|\}$ is a subspace of H; for we have $I - A^* A \geq 0$, so that $I - A^* A$ has a Hermitian square root, and consequently

$$\begin{aligned} \|Ax\| = \|x\| &\Longleftrightarrow ((I - A^* A)x,x) = 0 \\ &\Longleftrightarrow \|(I - A^* A)^{1/2}x\|^2 = 0 \\ &\Longleftrightarrow x \in \operatorname{Ker} (I - A^* A)^{1/2}. \end{aligned}$$

We may thus introduce subspaces V_i, $0 \leq i \leq n$, of H defined by

$$(1) \quad V_i = \begin{cases} H & \text{if } i = 0; \\ \{x \in H: \|A_1 A_2 \ldots A_i x\| = \|x\|\} & \text{if } i > 0. \end{cases}$$

It follows from the commutativity of the A's that $V_{i+1} \subseteq V_i$. We show that if $V_i \neq \{0\}$ then V_{i+1} is in fact a proper subspace of V_i. The restriction of A_i to V_i is an isometry for each i, and thus, for $x \in V_{i+1}$,

$$\|A_1 \ldots A_i(A_{i+1}x)\| = \|x\| = \|A_{i+1} x\|,$$

which implies that $A_{i+1}x \in V_i$. In other words, $A_{i+1}V_{i+1} \subseteq V_i$. If, for some i, $V_{i+1} = V_i \neq \{0\}$, then V_{i+1} is a non-trivial subspace of H, invariant with respect to A_{i+1}, on which A_{i+1} is isometric. It follows that A_{i+1} has an eigenvalue of unit modulus, contrary to hypothesis. It follows that the sequence $H \supseteq V_1 \supseteq V_2 \supseteq \ldots$ descends strictly until it reaches $\{0\}$, and hence $V_n = \{0\}$. This implies the desired conclusion $\|A_1 A_2 \ldots A_n\| < 1$.

The main result of the paper is in a kindred spirit.

Theorem 2. Let p be a polynomial of degree $k \geq 1$ having no two roots conjugate with respect to the unit circle and let g be the rational function given by the formula:

(2) $$g(z) = \frac{p(z)}{z^k p(1/\bar{z})^-}$$

<u>Let A be a linear operator on an n-dimensional Hilbert space</u> H <u>satisfying</u> $\|A\| \leq 1$, $|A|_\delta < 1$ <u>and rank</u> $(I - A^* A) = r$. <u>If</u> $rk \leq n$, <u>some root of</u> p <u>has modulus greater than one and</u> g(A) <u>is defined</u>, then $\|g(A)\| > 1$.

It should be emphasized that the hypothesis on the roots of p is supposed to include the condition that no root be conjugate to itself, i.e. be of unit modulus. The hypothesis thus implies that the expression (2) for g is in its lowest terms.

g(A) fails to be defined exactly when an eigenvalue of A coincides with a pole of q, or in other words, is conjugate to a root of p with respect to the unit circle. This can of course occur since the eigenvalues of A lie inside the unit circle while p is supposed to have a root outside it.

<u>Proof of Theorem 2</u>. We can suppose that p is monic. Let

$$p(z) = (z - \alpha_1)(z - \alpha_2) \ldots (z - \alpha_k),$$

and suppose that g(A) is defined. This means that $I - \bar{\alpha}_j A$ is non-singular, $1 \leq j \leq k$, so that we can introduce the operator A_j given by

(3) $$A_j = (A - \alpha_j I)\,(I - \bar{\alpha}_j A)^{-1}.$$

We have then

$$g(A) = A_1\, A_2 \ldots A_k,$$

and the A_j's all commute. We wish to introduce a descending sequence of subspaces of H, uch as in the proof of Theorem 1, but the definition (1) will no longer serve since we are dealing with A_j which need not satisfy $\|A_j\| \leq 1$. It turns out that a somewhat more complicated definition is appropriate.

For any subset σ of $\{1,\ldots,k\}$ let $A_\sigma = \prod_{j\in\sigma} A_j$ ($A_\emptyset$ being defined to be the identity operator), and let

$$(4)\quad V_i = \begin{cases} H & \text{if } i = 0 \\ \{x\in H: \|A_\sigma x\| = \|x\| \text{ for every } \sigma\subseteq\{1,\ldots,i\}\} & \text{if } 1\leq i\leq k. \end{cases}$$

Lemma. (i) V_i is a subspace of H of dimension at least $n - ri$, $0\leq i\leq k$.

(ii) V_i is a proper subspace of V_{i-1}, $1\leq i\leq n$.

(iii) If $x\in V_{i-1}\setminus V_i$, $1\leq i\leq n$, then $\|A_1A_2 \ldots A_ix\| \neq \|x\|$.

Proof of Lemma. It is an immediate consequence of the definition of V_i that the left hand side of the following identity is contained in the right hand side:

$$(5)\quad V_i = V_{i-1} \cap \{x\in H: \|A_ix\| = \|x\|\} \cap A_i^{-1} V_{i-1}.$$

Now consider any element x of the right hand side and any subset σ of $\{1,\ldots,i\}$. If $i \notin \sigma$ then $\|A_\sigma x\| = \|x\|$ since $x\in V_{i-1}$. Otherwise we may write $\sigma = \tau \cup \{i\}$, $\tau \subseteq \{1,\ldots,i - 1\}$, and (since the A_j's commute) $A_\sigma = A_\tau A_i$. Since $A_ix \subseteq V_{i-1}$ we have $\|A_\sigma x\| = \|A_\tau A_ix\| = \|A_ix\| = \|x\|$. Hence $x\in V_i$ and (5) is established.

Despite the fact that A_i need not be a contraction, it is still true that $\{x: \|A_ix\| = \|x\|\}$ is a subspace of H. This can be inferred from the Möbius identity

$$(6)\quad I - A_i^*A_i = (1 - \bar\alpha_i \alpha_i)(I - \alpha_iA^*)^{-1}(I - A^*A)(I - \bar\alpha_iA)^{-1}$$

which shows that $I - A_i^* A_i$ is either positive or negative semi-definite, depending on whether $|\alpha_i|< 1$ or $|\alpha_i|>1$, and hence that either $I - A_i^* A_i$ or $A_i^* A_i - I$ has a Hermitian square root. The standard argument given in the proof of Theorem 1 therefore applies. It is now easily seen with the help of an induction argument that V_i is a subspace of H for each i.

To prove the statement about dimensions we establish an alternative description of V_i. For $1 \leq i,\ j \leq k$ let $\nu(i,j)$ denote the number of indices ℓ in $\{1,2,\ldots,i\}$ for which $\alpha_\ell = \alpha_j$; thus $\nu(i,j)$ can be zero if $j > i$ but $\nu(i,j) \geq 1$ if $j \leq i$. We shall show that, for $1 \leq i \leq k$,

$$(7) \quad V_i = \{x \in H : (I - A^*A)(I - \bar{\alpha}_j A)^{-\nu} x = 0, \\ 1 \leq j \leq i,\ 1 \leq \nu \leq \nu(i,j)\}.$$

In the case $i = 1$, $\nu(1,1) = 1$ and, since $I - A_1^* A_1$ is semi-definite,

$$V_1 = \{x : ((I - A_1^* A_1)x, x) = 0\}$$
$$= \{x : (I - A_1^* A_1)x = 0\}$$
$$= \{x : (I - A^*A)(I - \bar{\alpha}_1 A)^{-1} x = 0\}$$

(the last step being by virtue of (6)), so that (7) holds.

Denote the right hand side of (7) by W_i and observe that

$$(8) \quad W_i = W_{i-1} \cap \operatorname{Ker} (I - A^*A)(I - \bar{\alpha}_i A)^{-\nu(i,i)}$$

so that, to establish (7) by induction, we must show that, for $x \in W_{i-1}$, $2 \leq i \leq k$, $\| A_i A_\sigma x \| = \| x \|$ for every subset σ of $\{1,\ldots,i-1\}$ if and only if $(I - A^*A)(I - \bar{\alpha}_i A)^{-\nu(i,i)} x = 0$.

Consider, then, $x \in W_{i-1} = V_{i-1}$ and $\sigma \subseteq \{1,\ldots,i-1\}$. By definition of V_{i-1}, $\| A_\sigma x \| = \| x \|$ and so $\| A_i A_\sigma x \| = \| x \|$ if and only if $\| A_i A_\sigma x \| = \| A_\sigma x \|$, or equivalently,

$$(A_\sigma^* (I - A_i^* A_i) A_\sigma x, x) = 0$$

or, by virtue of (6),

$$(9)\quad (A_{\sigma}^{*}(I - \alpha_i A^*)^{-1}(I - A^*A)(I - \bar{\alpha}_i A)^{-1}A_{\sigma} x,x) = 0.$$

Now $(I - \bar{\alpha}_i A)^{-1}A_{\sigma}$ is a rational function of A; its denominator is of degree one more than its numerator, and the factor $(I - \bar{\alpha}_j A)$ occurs in the denominator with an exponent which depends on the number of indices $\ell \in \sigma \cup \{i\}$ for which $\alpha_{\ell} = \alpha_j$. It will be seen that the partial fractions expression for $(I - \bar{\alpha}_i A)^{-1}A_{\sigma}$ can be written

$$(10)\quad (I - \bar{\alpha}_i A)^{-1}A_{\sigma} = \sum_{j=1}^{i} c_j(I - \bar{\alpha}_j A)^{-\nu(j,j)}$$

where $c_1,\dots,c_i$ are scalars depending on σ. Substituting (10) in (9) we deduce that $\|A_i A_{\sigma} x\| = \|x\|$ if and only if

$$(11)\quad \sum_{j,\ell=1}^{i} \bar{c}_j c_{\ell}\ ((I - \bar{\alpha}_j A^*)^{-\nu(j,j)} (I - A^*A)(I - \bar{\alpha}_{\ell} A)^{-\nu(\ell,\ell)}x,x) = 0.$$

Since $x \in W_{i-1}$ all the terms in (11) for which $j < i$ or $\ell < i$ vanish (this follows from the induction hypothesis), and (11) reduces to

$$|c_i|^2((I - \alpha_i A^*)^{-\nu(i,i)}(I - A^*A)(I - \bar{\alpha}_i A)^{-\nu(i,i)}x,x) = 0$$

and hence to

$$(12)\quad c_i(I - A^*A)(I - \bar{\alpha}_i A)^{-\nu(i,i)}x = 0.$$

It is now immediate that if $x \in W_i$ then $\|A_i A_{\sigma}' x\| = \|x\|$ for all $\sigma \subseteq \{1,\dots,i-1\}$, and hecne $x \in V_i$. To prove the converse, let $x \in V_i$ and choose σ to be $\{1,\dots,i-1\}$. Then in the resolution (10) of $(I - \bar{\alpha}_i A)^{-1}A_{\sigma}$ into partial fractions we have $c_i \neq 0$. To see this regard A as a scalar variable for the moment and notice that the left hand side of (10) has a pole of order $\nu(i,i)$ at $\lambda = 1/\bar{\alpha}_i$ (this depends on the hypothesis about the non-conjugacy of the α's, which ensures that no cancellation can take plase between numerator and denominator). The same must be true of the right hand side, and this can only be so if $c_i \neq 0$. Since $x \in V_i$, $\| A_i A_\sigma x \| = \| x \|$ and therefore (12) holds, and since $c_i \neq 0$ we have $(I - A^*A)(I - \bar{\alpha}_i A)^{-\nu(i,i)} x = 0$. It follows from (8) that $x \in V_i$. This completes the proof of the identity (7).

Statement (iii) of the Lemma also follows from the above argument. If $x \in V_{i-1}$ and $\| A_1 \dots A_i x \| = \| x \|$ then, as we have just shown, $(I - A^*A)(I - \bar{\alpha}_i A)^{-\nu(i,i)} x = 0$. But now that (7) is proved, we can rewrite (8)

(13) $\quad V_i = V_{i-1} \cap \operatorname{Ker} (I - A^*A)(I - \bar{\alpha}_i A)^{-\nu(i,i)}.$

Thus $x \in V_i$.

The relation (13) shows that the codimension of V_i in V_{i-1} is no greater than the rank of $I - A^*A$, which is r. Thus $\dim V_i \geq \dim V_{i-1} - r$, $1 \leq i \leq k$. Statement (i) follows at once.

The remaining assertion of the Lemma is justified much as in the proof of Theorem 1. By (i),

$$\dim V_{i-1} \geq n - r(i - 1) \geq n - rk + r \geq r.$$

Now r cannot be zero, else A would be unitary, contrary to the supposition $|A|_\sigma < 1$. Thus V_{i-1} is a non-null subspace, $1 \leq i \leq k$. From (5) we have $A_i V_i \subseteq V_{i-1}$, so that if $V_i = V_{i-1}$, V_i is a non-trivial invariant subspace for A_i and $A_i | V_i$ is an isometry. Thus A_i has an eigenvalue of unit modulus. However, referring to the definition (3) of A_i we perceive that the eigenvalues of A_i are of the form $(\lambda - \alpha_i)(1 - \bar{\alpha}_i \lambda)^{-1}$

where λ (being an eigenvalue of A) is not of unit modulus. It is well known (and easily checked) that $(\lambda-\alpha_i)(1-\bar{\alpha}_i\lambda)^{-1}$ cannot then be of unit modulus either (when $|\alpha_i| \neq 1$, as here). This contradiction shows that $V_i \neq V_{i-1}$.

We can now conclude the proof of Theorem 2. We can suppose that $|\alpha_k| > 1$. By (ii) of the Lemma there exists $x \in V_{k-1} \setminus V_k$. For such an x

$\| A_1 A_2 \ldots A_{k-1} x \| = \| x \|$ but, by (iii).

$$\| A_k A_1 A_2 \ldots A_{k-1} x \| \neq \| x \| .$$

Thus, if we write $A_1 A_2 \ldots A_{k-1} x = y$, we have $\| A_k y \| \neq \| y \|$. Now the Möbius identity (6) (with i = k) shows that $I - A_k^* A_k$ is negative semi-definite, which implies that $\| A_k u \| \geq \| u \|$ for every $u \in H$. It follows that $\| A_k y \| > \| y \| = \| x \|$. That is,

$$\| A_k A_1 A_2 \ldots A_{k-1} x \| > \| x \| .$$

Hence $\| A_1 A_2 \ldots A_k \| > 1$, as required.

<u>Corollary</u>. <u>Let</u> p <u>be a polynomial of degree</u> k <u>which is relatively prime to the polynomial</u> q <u>defined by</u>

$$q(z) = z^k \, p(1/z)^- .$$

<u>Let</u> A <u>be an</u> $n \times n$ <u>matrix such that</u> $|A|_\sigma < 1$ <u>and</u> $I - A^* A$ <u>is positive semi-definite and of rank</u> r. <u>If</u> $n \geq rk$, <u>the Hermitian form</u> $H(\cdot)$ <u>on</u> $\mathbb{C}^n$

$$H(x) = \| q(A)x \|^2 - \| p(A)x \|^2$$

<u>is positive semi-definite if and only if the roots of</u> p <u>all have modulus less than one</u>.

<u>Proof</u>. Suppose $H(\cdot)$ is positive semi-definite. Under these assumptions q(A) must be non-singular, for otherwise there is an eigenvalue λ of A such that $q(\lambda) = 0$. If we

take x to be a corresponding eigenvector we find that

$$H(x) = \| q(\lambda)x \|^2 - \| p(\lambda)x \|^2 = - | p(\lambda)|^2 \| x \|^2.$$

Since $H(x) \geq 0$, $p(\lambda) = 0$ and so p and q have a root in common, contrary to hypothesis. Another way of stating the assumption on $H(\cdot)$ is to say that

$q(A)^* q(A) - p(A)^* p(A) \geq 0$, or equivalently, $I - g(A)^* g(A) \geq 0$ where $g(A) = p(A)\, q(A)^{-1}$. This is in turn equivalent to $\| g(A) \| \leq 1$, and Theorem 2 shows that this can only hold if no root of p has absolute value greater than one. The assumption that p and q be relatively prime clearly implies that no root of p can lie on the unit circle, and hence all roots of p have modulus strictly less than one.

Suppose, conversely, that the roots of p all have modulus less than 1, and suppose further, for the moment, that q(A) is non-singular, so that we can form $g(A) = p(A)\, q(A)^{-1}$. In the notation of the proof of Theorem 2, $g(A) = A_1 A_2 \ldots$ $\ldots A_k$, while $\| A_j \| \leq 1$ for each j (this follows from (6)). Hence, by the submultiplicativity of the norm, $\| g(A) \| \leq 1$, and so $I - g(A)^* g(A) \geq 0$. Multiplying fore and aft by $q(A)^*$, $q(A)$, respectively, we infer that $q(A)^* q(A) - p(A)^* p(A) \geq 0$. Since the cone of positive semi-definite matrices is closed the restriction on q(A) can be removed by a simple continuity argument.

Making a suitable choice of A we derive a test akin to Schur's test for polynomials, but suffering from the disadvantage that we have to check p and q for common factors. If we stick to the case $r = 1$, as is a natural thing to do in practice, we can circumvent this awkward feature, and obtain a result which contains that of Schur.

Theorem 3. Let p *be a polynomial of degree* n, $n \geq 1$, *and let* A *be an* $n \times n$ *matrix such that* $|A|_{\sigma} < 1$ *and* $I - A^* A$ *is positive semi-definite and of rank one. The roots of* p *are of modulus less than one if and only if the Hermitian form* $H(\cdot)$ *on* $\underset{\sim}{C}^n$ *given by*

$$H(x) = \|q(A)x\|^2 - \|p(A)x\|^2$$

is positive definite, where q *is the polynomial defined by*

$$q(z) = z^n p(1/z)^-.$$

Proof. Suppose that the roots of p are of absolute value less than one. The roots of q must then all lie outside the unit circle, and since $|A|_{\sigma} < 1$, it follows that q(A) is non-singular. In the notation of the proof of Thoerem 2 (with k = n) we have $p(A)\, q(A)^{-1} = g(A) = A_1 A_2 \ldots A_n$. Now the A_j's commute and satisfy $\|A_j\| \leq 1$ (recall (6)); moreover, since $|A|_{\sigma} < 1$, it follows from properties of Möbius transformations that $|A_j|_{\sigma} < 1$. Hence, by Theorem 1, $\|g(A)\| < 1$. This is to say that $I - g(A)^* g(A)$ is positive definite, and hence that $q(A)^* q(A) - p(A)^* p(A)$ is positive definite, as required.

Conversely, let $H(\cdot)$ be positive definite. A fortiori $q(A)^* q(A)$ is positive definite, and hence q(A) is non-singular. Let the highest common factor of p and q be f: we wish to show that f = 1. Write $p = p_1 f$, $q = q_1 f$ where p_1, q_1 are relatively prime polynomials of degree k and $p_1(z) = (z - \alpha_1)\ldots(z - \alpha_k)$, and note that $q_1(z) = z^k p_1(m)^-$, as can readily be seen when p and q are written as products of linear factors. The identity

$q(A)^*q(A) = p(A)^*p(A) = q(A)^*[I - g_1(A)^*g_1(A)]\, q(A)$, where $g_1(A) = p_1(A)\, q_1(A)^{-1}$, and the fact that $q(A)$ is non-singular show that $I - g_1(A)^*g_1(A)$ is positive definite, and hence has kernel $\{0\}$. We have

$$\begin{aligned}\{0\} &= \operatorname{Ker}\, (I - g_1(A)^*g_1(A)) = \operatorname{Ker}\, (I - g_1(A)^*g_1(A))^{1/2} \\ &= \{x : \| g_1(A)x \| = \| x \|\} \\ &= \{x : \| A_1 \ldots A_k x \| = \| x \|\} \\ &\supseteq V_\nu ,\end{aligned}$$

where we are using once again the notation of Theorem 2 (with p replaced by p_1). Thus $V_k = \{0\}$. Since now $r = 1$, the Lemma (i) shows that $\dim V_k \geq n - k$, and hence $n = k$. Thus $p = p_1$ and $f = 1$. Now that we know that p and q are relatively prime we may apply the Corollary to Theorem 2 to deduce that the roots of p all lie in the open unit disc. This completes the proof of Theorem 3.

Theorem 3 gives us infinitely many ways of testing whether the roots of a polynomial have modulus less than one - one way for each choice of A satisfying $|A|_{\mathscr{G}} < 1$, $\| A \| \leq 1$ and rank $(I - A^*A) = 1$. The simplest such choice is $A = S$ where

$$S = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

(this is the matrix with respect to the natural basis of $\mathbb{C}^n$ of the shift operator S_n mentioned in the introduction). Then if

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_0,$$

so that

$$q(z) = \bar{a}_0 z^n + \bar{a}_1 z^{n-1} + \dots + \bar{a}_n,$$

we have

$$p(A) = \begin{bmatrix} a_0 & a_1 & \dots & a_{n-1} \\ 0 & a_0 & \dots & a_{n-2} \\ . & . & \dots & , \\ 0 & 0 & \dots & a_0 \end{bmatrix}$$

so that, if $x = x_0, \dots, x_{n-1})^T$,

$$p(A)x = (a_0x_0 + \dots + a_{n-1}x_{n-1}, a_0x_1 + \dots + a_{n-2}x_{n-1}, \dots$$
$$\dots, a_0x_{n-1})^T,$$

and so

$$\| p(A)x \|^2 = \sum_{j=0}^{n-1} | a_0x_j + a_1x_{j+1} + \dots + a_{n-j-1}x_{n-1} |^2 .$$

On performing a similar calculation for q we find that the Hermitian form H( ) is in this case

$$H(x) = \sum_{j=0}^{n-1} | a_nx_j + a_{n-1}x_{j+1} + \dots + a_{j+1}x_{n-1}|^2$$
$$- \sum_{j=0}^{n-1} | a_0x_j + a_1x_{j+1} + \dots + a_{n-j-1}x_{n-1} |^2 .$$

This yields precisely Schur's criterion [5, § 13].

It is conceivable that there might on occasion be something to be gained by making a different choice of A. Now the class $\mathcal{E}$ of matrices A satisfying the conditions we need ($\| A \| \leq 1$, $|A|_\sigma < 1$, rank $(I - A^*A) = 1$) has arisen in some closely related investigations of the author and V. Pták, and has been shown to have some very interesting properties. For example, $\mathcal{E}$ consists of all matrices unitarily equivalent to a matrix of the form $(AS + B)(CS + D)^{-1}$ where A,B,C,D are

$n \times n$ matrices such that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^* \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$$

(see [4, Theorem 1]). This description of $\mathcal{E}$ enables us to write down quite a large class of members of $\mathcal{E}$ explicitly ([4, § 2]): $\mathcal{E}$ contains all matrices of the form

$$A = \begin{bmatrix} p_1 & s_1 s_2 & -s_1 \overline{p}_1 s_2 & \dots & (-1)^n s_1 \overline{p}_2 \dots \overline{p}_{n-1} s_n \\ 0 & p_2 & s_2 s_3 & \dots & (-1)^{n-1} s_2 \overline{p}_3 \dots \overline{p}_{n-1} s_n \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & p_n \end{bmatrix}$$

where $p_1, \dots, p_n$ are arbitrary complex numbers of modulus less than 1, and $s_j = (1 - p_j \overline{p}_j)^{1/2}$, $1 \leq j \leq n$.

Acknowledgement. I should like to express my gratitude to the Mathematical Institute of the Czechoslovak Academy of Sciences, and in particular, to the Department of Functional Analysis, for their hospitality and financial assistance during the period when this work was carried out.

References

[1] R. NEVANLINNA: Über beschränkte Funktionen, die in gegebenen Punkten vorgeschriebene Werte annehmen, Ann. Acad. Sci. Fennicae 13, No 1(1919).

[2] V. PTÁK: Norms and the spectral radius of matrices, Czechosl. Math. J. 87(1962), 553-557.

[3] V. PTÁK: Isometric parts of operators and the critical exponent, Časopis pro pěstování matematiky 101 (1976), 383-388.

[4] V. PTÁK and N.J. YOUNG: Functions of operators and the spectral radius (to appear).

[5] J. SCHUR: Uber Potenzreihen, die im Innern des Einheitskreises beschränkt sind, J. für Mathematik 147 (1917), 205-232 and 148(1918), 122-145.

[6] N.J. YOUNG: Matrices which maximise any analytic function, to appear.

Matematický ústav ČSAV
Žitná 25, 11567 Praha 1
Československo
(till 31st December,1978)

Department of Mathematics
The University Glasgow
Scotland
(from 1st January,1979)

(Oblatum 26.7. 1978)