Vlastimil Pták

Isometries in H^2, generating functions and extremal problems

Časopis pro pěstování matematiky, Vol. 110 (1985), No. 1, 33--57

Persistent URL: `http://dml.cz/dmlcz/118220`

Časopis pro pěstování matematiky, roč. 110 (1985), Praha

ISOMETRIES IN H^2, GENERATING FUNCTIONS AND EXTREMAL PROBLEMS

VLASTIMIL PTÁK, Praha

(Received March 8, 1983)

INTRODUCTION

Suppose T is an isometry on a Hilbert space H so that $T^*T = 1$. Consider the operator T^*. Since $T^*(1 - TT^*) = 0$ the range of $1 - TT^*$ is contained in the subspace Ker T^*. On the other hand, if $T^*x = 0$, then $x = x - TT^*x = (1 - TT^*)x$. We see thus that Ker T^* equals the range of $1 - TT^*$.

This, of course, is true for any partial isometry T: the operator T^*T is the projection onto the initial space of T and TT^* is the projection onto the final space of T. The complement of the final space, the range of $1 - TT^*$, is thus equal to Ker TT^* but this is nothing more than Ker T^*.

In particular T^* is injective if and only if T is unitary. Thus for a nonunitary isometry T the space

$$\operatorname{Ker} T^* = \operatorname{Range}(1 - TT^*) = H \ominus TH$$

is nontrivial. It is customary to call it the wandering subspace for T the name being justified by the fact that

$$T^p x \perp T^q y$$

for any pair $0 \leqq p < q$ and $x, y \in \operatorname{Ker} T^*$.

This fact makes it possible to define an operator $R(T)$ which intertwines T and a shift operator V on $H^2(\operatorname{Ker} T^*)$

$$R(T)\, V = T\, R(T)\,.$$

The operator V is a unilateral shift of dimension $d = \dim \operatorname{Ker} T^*$.

Similar ideas are frequently used in the theory of unitary dilations of contractions treated in the book [11] of B. Sz. Nagy and C. Foias where the reader can find deeper information concerning this matter.

An important particular case of an isometry in H^2 is the operator of multiplication by a given inner function φ, the mapping T which assigns to each $f \in H^2$ the product

$\varphi(z) f(z)$. If φ is a Blaschke product of length n then the corresponding wandering subspace Ker T^* has dimension n.

In the present note we intend to investigate somewhat closer the case of an isometry generated by a single Möbius function

$$\varphi(z) = \frac{z - \alpha}{1 - \alpha^* z} .$$

We establish several interesting relations between the function φ and the corresponding intertwining operator $R(\varphi)$. As an application, we use these relations to find an explicit expression for the sequence of vectors obtained by the Gram-Schmidt orthonormalization process from a given sequence in H^2. This presents yet another method of describing the matrix of the operator $S \mid \text{Ker}\,(S - \alpha)^n$ with respect to an orthonormal basis.

Here S is the (backward) shift operator on H^2 defined by

$$(Sf)(z) = \frac{1}{z}(f(z) - f(0))$$

and α is some number less than one in modulus.

The importance of the problem of finding a concrete representation for this operator lies in the fact that it was shown in [2, 3] that this operator realizes the maximum of $|A^n|$ under the constraints that A is a contraction whose spectral radius does not exceed r if r is a positive number less than one (we take $|\alpha| = r$).

It is interesting to note that – taken quite formally – the method to be described here was essentially the first method by means of which the concrete representation was obtained.

In the author's paper [2, 3] the extremal operator was identified as $S \mid \text{Ker}\,(S - r)^n$.

For a further study of this operator, it was obviously necessary to have its matrix with respect to an orthonormal basis.

In a series of discussions of M. Fiedler and the author such a matrix was obtained using the method of generating functions in a purely heuristic manner without formally justifiable proofs. Because of the unorthodox manner in which the result has been obtained it was never published. The first published proof is that of N. J. Young [12] which is based on complicated algebraic manipulations. For further developments, see the survey [8]. When the author became acquainted with the theory of Nagy-Foias it became obvious that the ideas of this theory could be used to justify the formal computations with generating functions. Surprisingly enough this turns out to be less straightforward than it might seem at first glance. This note represents a somewhat improved version of the author's notes on this matter: it seems that the properties of the isometry generated by a Möbius function are interesting on their own and that the application which we present in chapter three contributes to the rehabilitation of generating functions.

1. MATRICES AND GENERATING FUNCTIONS

We denote by l^2 the Hilbert space of all sequences

$$\{x_0, x_1, x_2, \ldots\}$$

of complex numbers such that $\Sigma|x_j|^2 < \infty$ with the usual algebraic operations and scalar product; the open unit disc and the unit circle will be denoted by D and T respectively.

We shall frequently identify the element $a \in l^2$ with the corresponding element $\Sigma a_n z^n \in L^2(T)$ or with the corresponding holomorphic function in the Hardy space H^2. The standard basis of l^2 will be denoted by $e_0, e_1, e_2, \ldots, P_+$ will be the orthogonal projection of $L^2(T)$ onto H^2, the closed linear span of the functions $1, z, z^2, \ldots$. We write V for the operator of multiplication by z on $L^2(T)$ and S for the backward shift on H^2, so that $S = (V| H^2)^*$. If g is an element of H^2 we denote by $\tilde{g}$ the function $\tilde{g}(z) = g(z^*)^*$. For every $x \in D$ let us denote by $p(x)$ the element of l^2

$$p(x) = \{1, x, x^2, \ldots\}.$$

We shall denote by $e(x)$ the element $p(x^*)$; it will be called the evaluation functional corresponding to x since it may be used to obtain the value of the function in H^2 corresponding to an element $f \in l^2$ as follows

$$f(x) = (f, e(x)), \quad \text{for all} \quad x \in D.$$

A doubly infinite array of complex numbers b_{ij}, $i, j = 0, 1, 2, \ldots$, will be called an operator matrix if the following two conditions are satisfied

1° for each i the sequence $b_i = \{b_{ij}\}_j$ belongs to l^2

2° for each $x \in l^2$ the sequence

$$\{\Sigma b_{0k}x_k, \ \Sigma b_{1k}x_k, \ \Sigma b_{2k}x_k, \ \ldots\} \quad \text{belongs to } l^2.$$

The bilinear form corresponding to an operator matrix B is defined as the expression

$$(B\, p(x), p(y)) = \Sigma b_{jk} x^k y^{*j}$$

for x and y in D. It is not difficult to see that an operator matrix is uniquely determined by its bilinear form. If we assign to the operator matrix B the sequence u_j of vectors in l^2

$$u_j = \{b_{jk}^*\}_k$$

then the condition 2° may be reformulated as follows

3° for each $x \in l^2$ the sequence

$$F(x) = \{(x, u_0), (x, u_1), (x, u_2), \ldots\}$$

belongs to l^2.

Let us show now that F is a bounded linear operator in l^2: this is a standard closed graph argument. Indeed, suppose that $x_n \to 0$ and $F(x_n) \to m$. Since $(F(x), e_k) = (x, u_k)$ for all x, we have

$$(m, e_k) = \lim_n (F(x_n), e_k) = \lim_n (x_n, u_k) = 0$$

so that m is zero and F continuous.

We shall use this fact to show that if B is an operator matrix then B^{T} is an operator matrix as well. Indeed, we observe that

$$(Fe_k, e_j) = (e_k, u_j) = b_{jk}$$

so that the column $\{b_{jk}\}_j$ consists of the Fourier coefficients of Fe_k and thus belongs to l^2. Given any $y \in l^2$, we may thus form the sequence

$$\{\Sigma b_{j0} y_j,\ \Sigma b_{j1} y_j,\ \Sigma b_{j2} y_j,\ \ldots\} = \{(Fe_0, z),\ (Fe_1, z),\ (Fe_2, z),\ \ldots\}$$

which is nothing more than the sequence of conjugate Fourier coefficients of F^*z if z stands for the element $\{\bar{y}_0, \bar{y}_1, \bar{y}_2, \ldots\}$.

Given a bounded linear operator A on H^2 it is easy to see that the double sequence

$$a_{jk} = (Ae_k, e_j)$$

is an operator matrix. We shall call it the matrix of A or the matrix corresponding to A and denote it by $\mathcal{M}(A)$. If we use the terminology and notation just introduced we see that the operator F discussed above is defined in such a manner that its matrix is exactly B. The function $A(s, t)$ defined on $D \times D$ by the formula

$$A(s, t) = (A\, e(t), e(s))$$

will be called the kernel of A.

The relations among the notions just introduced are described by the following propositions

(1,1) *For any operator matrix B the array B^{T} is an operator matrix as well and*

$$(B\, p(y), p(x)) = (B^{\mathsf{T}}\, e(x), e(y))$$

$$(B\, p(v^*), p(u^*)) = (B^{\mathsf{T}}\, p(u), p(v)).$$

Proof. $(B^{\mathsf{T}}\, e(x), e(y)) = \Sigma(B^{\mathsf{T}})_{rs}\, x^{*s} y^r = \Sigma b_{sr} x^{*s} y^r = (B\, p(y), p(x))$.

(1,2) *Given a bounded linear operator A on H^2 the matrix $\mathcal{M}(A)$ is characterized by the relation*

$$(A\, p(x), p(y)) = (\mathcal{M}(A)\, p(x), p(y)).$$

Proof. The reader will understand that the pedantically rigorous statement of this lemma would require introducing a notation distinguishing between $p(x)$ taken as a sequence in l^2 and as an element of H^2 and will be able to give each such statement its correct interpretation. The jk-th element of $\mathcal{M}(A)$ is given by

$$\mathcal{M}(A)_{jk} = (Ae_k, e_j)$$

so that

$$(A\, p(x), p(y)) = (A\, \Sigma x^k e_k, \Sigma y^j e_j) = \Sigma x^k \mathcal{M}(A)_{jk}\, y^{*j} = (\mathcal{M}(A)\, p(x), p(y)).$$

(1,3) *Given a bounded linear operator A on H^2 then the kernel corresponding to A is the bilinear form corresponding to $\mathcal{M}(A)^{\mathsf{T}}$.*

Proof. Using (1,2) and (1,1) we obtain

$$A(s, t) = (A\, e(t), e(s)) = (\mathcal{M}(A)\, e(t), e(s)) = (\mathcal{M}(A^{\mathsf{T}})\, p(s), p(t)).$$

(1,4) *The kernel of a bounded linear operator A possesses the following property*

$$(Af)(s) = \frac{1}{2\pi i}\int A(s, z) f(z) \frac{dz}{z}$$

for any $f \in H^2$, the integral being taken counterlockwise on the unit circle.

We shall not give a proof of this fact which is only included for the sake of completeness. The statement itself requires some amplification. Indeed, we have defined $A(s, z)$ as $(A\, e(z), e(s))$ for s and z less than 1 in modulus, the evaluation functionals $e(z)$ are meaningless for $|z| = 1$. Nevertheless, it may be shown that $A(\cdot, t) \in H^2$ and $\overline{A(s, \cdot)} \in H^2$ for all $s, t \in D$. As functions in H^2 they have well defined boundary values and it is in this sense that the above integral is to be taken.

2. ISOMETRIES CORRESPONDING TO A MÖBIUS FUNCTION

In this section we intend to study a class of matrices which correspond in a natural manner to isometries in H^2 generated by a Möbius function.

Let us denote by $\mathcal{M}$ the set of all two by two matrices A which satisfy

$$A^*QA = Q$$

where Q is the matrix

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

It is easy to see that, for a matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

the inclusion $A \in \mathcal{M}$ is equivalent to the following three relations to be satisfied for the entries a, b, c, d:

$$|a|^2 - |c|^2 = 1$$
$$|d|^2 - |b|^2 = 1$$
$$a^*b - c^*d = 0 ;$$

if these conditions are satisfied, we clearly have $|a| > 0$, $|d| > 0$ and it is not difficult to show that

$$\frac{c}{d} = \left(\frac{b}{a}\right)^* .$$

Indeed, multiplying the equation

$$a^*b = c^*d$$

by ca we obtain $cb|a|^2 = |c|^2 ad$ so that

$$\frac{c}{d} \cdot \frac{b}{a} = \frac{|c|^2}{|a|^2} \geqq 0 .$$

The equation $a^*b = c^*d$ implies $|a|\,|b| = |c|\,|d|$; since

$$|a|^2\,|b|^2 = (1 + |c|^2)\,|b|^2$$
$$|c|^2\,|d|^2 = |c|^2\,(1 + |b|^2)$$

this implies $|b| = |c|$.

Furthermore

$$|a|^2 = 1 + |c|^2 = 1 + |b|^2 = |d|^2$$

so that $|a| = |d|$. The numbers c/d and b/a have thus the same modulus so that they have to be complex conjugate.

It follows from $|a| = |d| > 0$ that the quotient $\varepsilon = a/d$ has modulus 1. If we denote by α the quotient $- b/a$ we have $c/d = \alpha^*$; at the same time $|a|^2 - |c|^2 = 1$ and $|b| = |c|$. Thus $|a|^2 - |b|^2 = 1$ whence $|a|^2 = |b|^2 + 1 > |b|^2$ so that $|\alpha| = (1 - |a|^{-2})^{1/2} < 1$.

For $x \in D$ we have thus

$$\frac{ax + b}{cx + d} = \varepsilon \frac{x - \alpha}{1 - \alpha^* x} .$$

In the rest of this section we shall prove that, for each $A \in \mathcal{M}$, there exists an operator matrix $m(A)$ whose bilinear form equals

$$(-bx + d - y^*(ax - c))^{-1} = (cy^* + d - x(ay^* + b))^{-1} .$$

We intend to describe now an operator on H^2 whose matrix is $m(A)$; in other words we intend to construct an operator $R(A)$ such that

$$\mathcal{M}(R(A)) = m(A).$$

The construction of this operator is a particular case of a more general set-up. It will therefore be convenient to begin by recalling some known facts about isometries and wandering spaces.

The wandering space of an isometry T on a Hilbert space H is defined as the subspace

$$\operatorname{Ker} T^* = \operatorname{Range}(1 - TT^*) = H \ominus TH.$$

If we denote by $H^2(\operatorname{Ker} T^*)$ the Hilbert space of all sequences $h_0, h_1, \ldots$ with $h_j \in \operatorname{Ker} T^*$, $\Sigma|h_j|^2 < \infty$ and by W the forward shift operator on $H^2(\operatorname{Ker} T^*)$

$$W(h_0, h_1, \ldots) = (0, h_0, h_1, \ldots)$$

then there exists a natural intertwining operator R from $H^2(\operatorname{Ker} T^*)$ into H^2 such that

$$RW = TR.$$

Indeed, if we write D for $\operatorname{Ker} T^*$ and if we identify $H^2(\operatorname{Ker} T^*)$ with $D \otimes H^2$ then the relation

$$R(d \otimes e_k) = T^k d$$

defines a mapping of $D \otimes H^2$ into H such that

$$TR(d \otimes e_k) = T^{k+1}d = R(d \otimes e_{k+1}) = RW(d \otimes e_k)$$

for every $k = 0, 1, 2, \ldots$ whence

$$TR = RW.$$

It follows from this relation that the range of R is invariant with respect to T. It is interesting to note that, in fact, it reduces T. It suffices to observe that the orthogonal complement of the range of R is invariant with respect to T as well; indeed, it is easy to identify it as the intersection $\bigcap T^n H$.

To see that, consider an $n \geqq 1$ and an arbitrary $x \in H$. Then

$$(T^n x, T^{n-1}D) = (Tx, D) = (x, T^*D) = 0$$

so that $\operatorname{Range} T^n \subset (T^{n-1}D)^\perp$ and

$$\bigcap_{n \geqq 1} (\operatorname{Range} T^n) \subset \bigcap_{n \geqq 1} (T^{n-1}D)^\perp \subset (\operatorname{Range} R)^\perp.$$

On the other hand, if $x \perp \operatorname{Range} R$ then

$$x \in D^\perp \cap (TD)^\perp \cap (T^2D)^\perp \cap \ldots.$$

Now $x \in D^\perp$ implies $x = Ty$ for some y; since $(y, D) = (Ty, TD) = (x, TD) = 0$, the element y itself lies in the range of T so that $y = Tz$, $x = T^2 z$. In a similar manner,

$(z, D) = (T^2 z, T^2 D) = (x, T^2 D) = 0$. so that $z = Tv$ for some v and so on; this shows that $x \in$ Range T^n for every $n = 1, 2, \ldots$. This proves that

$$(\text{Range } R)^{\perp} = \bigcap_n \text{Range } T^n .$$

Consider now a fixed element $h \in D$, $|h| = 1$ and denote by T_1 the restriction of T to the closed linear span M of $h, Th, T^2 h, \ldots$. Then

$R(T_1)$ is a mapping of $H^2(\text{Ker } T_1^*) \sim H^2$ into M such that

$$R(T_1)\, V = T_1\, R(T_1)$$

where V is the forward shift on $H^2(\text{Ker } T_1^*)$. If f is any polynomial then

$$R(T_1) f(V) = f(T)\, R(T_1) .$$

The isometry $R(T_1)$ transforms thus the polynomial $f \in H^2$ into the element $f(T)\, h$. It is obvious that this correspondence $f \to f(T)\, h$ may be extended to a larger class of functions, at least to functions holomorphic in a neighbourhood of the unit disc. In particular, the elements $e(x)$ for $x \in D$ are such functions. Let us compute

$$(R(T_1)\, e(x), e(z)) = (R(T)\, \Sigma x^{*k} e_k, e(z)) =$$

$$= (\Sigma x^{*k} R(T)\, e_k, e(z)) = (\Sigma x^{*k} T^k h, e(z)) = ((1 - x^* T)^{-1} h)(z) .$$

An important particular case of an isometry is the multiplication by an inner function on H^2; if we denote the multiplication operator by $M(\varphi)$ then $M(\varphi)^* M(\varphi) = 1$. Since $M(\varphi)^* = P_+ M(\bar{\varphi})$ the final space of $M(\varphi)$ will be $M(\varphi)\, M(\varphi)^* = M(\varphi)\, P_+ M(\bar{\varphi})$ so that its orthogonal complement will be

$$1 - M(\varphi)\, M(\varphi)^* = 1 - M(\varphi)\, P_+ M(\bar{\varphi}) =$$

$$= M(\varphi)\, M(\bar{\varphi}) - M(\varphi)\, P_+ M(\bar{\varphi}) =$$

$$= M(\varphi)\, (1 - P_+)\, M(\bar{\varphi}) = M(\varphi)\, P_- M(\bar{\varphi}) .$$

For the wandering subspace we have the following alternative descriptions.

$$\text{Ker } M(\varphi)^* = \text{Ker } \varphi(V)^* = \text{Ker } \tilde{\varphi}(S) = H^2 \ominus \varphi H_2 .$$

If φ is a Blaschke product of length d then the dimension of the wandering subspace Ker $M(\varphi)^*$ is exactly d.

If T is the isometry $M(\varphi)$ with

$$\varphi(z) = \frac{z - \alpha}{1 - \alpha^* z}, \quad (|\alpha| < 1)$$

then Ker T^* is the one-dimensional subspace consisting of all scalar multiples of the function

$$\frac{1}{1 - \alpha^* z} .$$

This may be easily seen as follows: since $T^*f = P_+ \overline{\varphi(z)} f(z)$ the condition $T^*f = 0$ may be successively reformulated in the following equivalent forms

$$\overline{\varphi(z)} f(z) \in H^2_-$$

$$\bar{z}\, \overline{\varphi(\bar{z})} f(\bar{z}) \in H^2$$

$$\frac{1 - \bar{\alpha}\bar{z}}{1 - \alpha z} f(\bar{z}) \in H^2$$

and this is possible if and only if f is a scalar multiple of the function

$$\frac{1}{1 - \bar{\alpha} z}.$$

Let us turn our attention to the operator $R(M(\varphi))$; according to what has been said above it is defined by the relations

$$R(M(\varphi))\, e_k = \left(\frac{z - \alpha}{1 - \alpha^* z}\right)^k \frac{(1 - |\alpha|^2)^{1/2}}{1 - \alpha^* z}.$$

Now denote by M the operator of multiplication by

$$\frac{az + b}{cz + d} \quad \text{if} \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}.$$

We have seen that

$$(R(M)\, p(x), p(y)) = (R(M)\, e(x^*), e(y^*)) =$$

$$= ((1 - xM)^{-1}\, h)\,(y^*) = \frac{1}{1 - x\dfrac{ay^* + b}{cy^* + d}}\, h(y^*).$$

Now

$$h(z) = \frac{1}{d}\, \frac{1}{1 + \dfrac{c}{d} z} \cdot |d| \left(1 - \left|\frac{c}{d}\right|^2\right)^{1/2} =$$

$$= \frac{1}{d + cz} (|d|^2 - |c|^2)^{1/2} = \frac{1}{d + cz} (|d|^2 - |b|^2)^{1/2} = \frac{1}{cz + d}.$$

It follows that

$$(R(M)\, p(x), p(y)) = \frac{1}{cy^* + d - x(ay^* + b)} = (m(A)\, p(x), p(y))$$

whence

$$\mathcal{M}(R(M)) = m(A).$$

To obtain the kernel $K(s, t)$ of the operator $R(M)$ it suffices to take the bilinear form

$$(R(M)\, p(x), p(y)) \quad \text{for the values} \quad x = t^*\,, \quad y = s^*\,.$$

Thus $K(s, t) = (cs + d - t^*(as + b))^{-1}$.

It is interesting to verify this directly.

Indeed, consider an arbitrary polynomial f and let us compute the integral

$$\frac{1}{2\pi i}\int \frac{1}{cx + d - y^*(ax + b)} f(y) \frac{\mathrm{d}y}{y}$$

taken counterlockwise on the unit circle; it may be rewritten in the form

$$\frac{1}{cx + d}\,\frac{1}{2\pi i}\int \frac{f(y)}{y - \dfrac{ax + b}{cx + d}}\,\mathrm{d}y$$

which, by the Cauchy integral formula, is nothing more than

$$\frac{1}{cx + d} f\left(\frac{ax + b}{cx + d}\right).$$

Summing up, we have assigned to every two by two matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

such that $A \in \mathscr{M}$ an operator R in H^2 such that

$$\mathscr{M}(R) = m(A)\,.$$

This relation possesses the following equivalent reformulations

$$(R\, p(x), p(y)) = (m(A)\, p(x), p(y)) =$$

$$= (-bx + d - y^*(ax - c))^{-1} = (cy^* + d - x(ay^* + b))^{-1}\,,$$

$$R\, p(x) = \frac{1}{-bx + d} \quad p\left(\frac{ax - c}{-bx + d}\right),$$

$$(Rf)(z) = \frac{1}{cz + d} \quad f\left(\frac{az + b}{cz + d}\right).$$

3. SUBSTITUTION OPERATORS

We have seen that there exist naturally defined operators in the space H^2 whose generating functions are derived from two by two matrices corresponding to Möbius functions. It is natural to ask whether this correspondence could not be extended to more general matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

It is useful to note that, under the assumption $|c| < |d|$, the expression $(az + b)$. . $(cz + d)^{-1}$ reduces to a Möbius function as follows: for $|z| < 1$, we have

$$\frac{az + b}{cz + d} = \frac{bd^* - ac^*}{|d|^2 - |c|^2} + \frac{ad - bc}{|d|^2 - |c|^2} \frac{d^*}{d} \frac{z + \alpha}{1 + \alpha^* z}$$

with $\alpha = (cd^{-1})^*$

For the application that we have in mind it will be necessary to use matrices more general than those corresponding to Möbius functions.

We shall see that a parallel theory may be developed even for this more general case.

We do not intend to embark upon a general study of substitution operators of this type since we have a definite purpose in mind — to show how they can be used to obtain a concrete representation for our extremal operators. Let us remark that the extension to be considered in the present paragraph consists essentially in replacing the equality $A^*QA = Q$ by the inequality $A^*QA \leqq Q$.

We shall not pursue this idea further; we might return to it in another communication We shall proceed in a completely elementary manner; we begin by showing that, for each matrix A which yields a bounded substitution operator, the matrix QA^TQ possesses the same properties. A proposition which we prove later will present this fact in another light connecting it with the properties of the adjoint to a substitution operator.

(3,1) *Suppose that* $|c| < |d|$ *and that*

$$\frac{az + b}{cz + d} \in D$$

for each $z \in D$. *Then* $|b| < |d|$ *and*

$$\frac{az - c}{-bz + d} \in D$$

for all $z \in D$.

Proof. Since $|c| < |d|$ the quotient

$$\frac{az + b}{cz + d}$$

is defined for all z with $|z| \leqq 1$.

We shall denote it by φ. According to our assumption $\varphi(D) \subset D$. In particular $\varphi(0) \in D$ and this means that $|b| < |d|$. Now suppose that $|z_0| < 1$ and

$$\left|\frac{az_0 - c}{-bz_0 + d}\right| \geqq 1 .$$

It follows that the number

$$y = \frac{-bz_0 + d}{az_0 - c}$$

has absolute value $|y| \leqq 1$. Now $(ay + b)\, z_0 = cy + d$. Since $|y| \leqq 1$ and $|c| < |d|$ we have $z_0 \neq 0$. Thus

$$\varphi(y) = \frac{1}{z_0}$$

so that $|\varphi(y)| > 1$. This is a contradiction if $|y| < 1$. If $|y| = 1$, there exists a positive $r < 1$ such that $|\varphi(ry)| > 1$, a contradiction again.

(3,2) *Suppose T is a bounded linear operator on H^2 such that*

$$(Te_k, e(z)) = \frac{(az + b)^k}{(cz + d)^{k+1}}$$

for every $z \in D$.
Then

$$(T^*e_k, e(z)) = \frac{(a^*z - c^*)^k}{(-b^*z + d^*)^{k+1}}$$

for all $z \in D$.

Proof. Given $y \in D$,

$$(T^*e_r, e(y)) = (e_r, T\Sigma y^{*k} e_k) = \left(e_r, \Sigma y^{*k} \frac{(az + b)^k}{(cz + d)^{k+1}}\right) =$$

$$= \left(e_r, \frac{1}{cz + d - y^*(az + b)}\right) = \left(e_r, \frac{1}{d - by^* - z(ay^* - c)}\right) =$$

$$= \left(\frac{1}{-by^* + d}\right)^* \left(e_r, \Sigma z^k \left(\frac{ay^* - c}{-by^* + d}\right)^k\right) = \left(\frac{(ay^* - c)^r}{(-by^* + d)^{r+1}}\right)^* .$$

Our first task will be to describe those matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

for which there exists a bounded linear operator T on H^2 such that

$$(14) \qquad (Tp(x), p(y)) = \frac{1}{-bx + d - y^*(ax - c)}$$

for all $|x| < 1$ and $|y| < 1$. We shall denote this class of matrices by $\mathcal{M}_0$. For brevity, we shall write $T = m(A)$ if (14) is satisfied.

Consider first the function appearing in the denominator: it may be expressed in two different ways

$$F(x, y) = (-bx + d - y^*(ax - c)) = ((cy^* + d) - x(ay^* + b)).$$

Let us state first some conditions which are necessary in order that F be different from zero for all $|x| < 1$ and $|y| < 1$.

(3,3) *If $F(x, y)$ is different from zero for all $x \in D$ and $y \in D$ then*

1° $d \neq 0$,

2° $|at - c| \leqq |-bt + d|$ *and* $|at + b| \leqq |ct \times d|$ *for all* $t \in D$,

3° $|b| \leqq |d|$ *and* $|c| \leqq |d|$.

Proof. Suppose that F is different from zero for all $x \in D$ and $y \in D$. The first assertion is a consequence of the equality $d = F(0, 0)$. Suppose that $|ax - c| > |-bx + d|$ for some $x \in D$; then

$$w = \frac{-bx + d}{ax - c} \in D$$

and $F(x, w^*) = 0$ which is a contradiction. If $|ax + b| > |cx + d|$ for some $x \in D$ then

$$v = \frac{cx + d}{ax + b} \in D$$

and $F(v, x^*) = 0$, a contradiction again. This proves the second assertion. The third assertion follows from the second one upon taking $x = 0$. The proof is complete.

Two observations are immediate: first, it follows from the preceding lemma that matrices belonging to $\mathcal{M}_0$ have to satisfy the three conditions from Proposition (3,3), second, $\mathcal{M}_0$ contains $\mathcal{M}$. It turns out that operators satisfying (14) are substitution operators and that there are two kinds of these according to whether A is singular or not.

The situation is completely described in the following two propositions.

(3,4) *Suppose A satisfies the conditions from Proposition* (3,3). *Then these are equivalent.*

1° $A \in \mathcal{M}_0$ *and the operator* $m(A)$ *is one-dimensional,*

2° *both* $|b|$ *and* $|c|$ *are less than* $|d|$ *and* $\det A = 0$,

3° $A \in \mathcal{M}_0$ *and* A *is singular.*

In this case the operator $T = m(A)$ may be described as follows

$$(Tf)(z) = \frac{1}{cz + d} f\left(\frac{b}{d}\right),$$

$$Tp(x) = \frac{1}{-bx + d} p\left(-\frac{c}{d}\right).$$

Furthermore, for a matrix A satisfying the three conditions from Proposition (3,3) *the following four conditions are equivalent*

1° $A \in \mathscr{M}_0$ *and $m(A)$ is not one-dimensional,*

2° $A \in \mathscr{M}_0$ *and $m(A)$ is infinite dimensional,*

3° A *is nonsingular,*

4° A *is nonsingular and* $\left|\dfrac{az - c}{-bz + d}\right| < 1$ *for* $|z| < 1$.

In this case the operator $T = m(A)$ may be described as follows

$$(Tf)(z) = \frac{1}{cz + d} f\left(\frac{az + b}{cz + d}\right) \quad for \quad f \in H^2, \quad z \in D.$$

$$Tp(x) = \frac{1}{-bx + d} p\left(\frac{ax - c}{-bx + d}\right) \quad for \quad x \in D.$$

Proof. Suppose that $T = m(A)$ and $T = hg^*$. Then

$$(-bx + d - y^*(ax - c))^{-1} = (Tp(x), p(y)) = p(y)^* hg^* p(x) = h(y^*) \tilde{g}(x);$$

in particular

$$h(0) \tilde{g}(x) = (-bx + d)^{-1}$$

whence

$$h(y^*) = (\tilde{g}(x))^{-1} h(y^*) \tilde{g}(x) =$$

$$= h(0)(-bx + d)(-bx + d - y^*(ax - c))^{-1} = h(0)\left(1 - y^* \frac{ax - c}{-bx + d}\right)^{-1}.$$

It follows that $(ax - c)/(-bx + d)$ is a constant ξ independent of x; since $h \in H^2$ we must have $|\xi| < 1$. Since $\tilde{g}$ is a multiple of $(-bx + d)^{-1}$ the inclusion $\tilde{g} \in H^2$ implies $|b| < |d|$. Now $ax - c = \xi(-bx + d)$ so that $ad - bc = 0$.

Consider first the case $b = 0$. In this case $a = 0$ as well and

$$h(y^*) \tilde{g}(x) = (cy^* + d)^{-1}.$$

Thus g is a constant and

$$h(z) g(0)^* = (cz + d)^{-1}$$

whence

$$Tf = (f, g)\, h(z) = f(0)\, g(0)^*\, h(z) = f(0) \frac{1}{cz + d}.$$

If $b \neq 0$ then $\xi = -c/d = -a/b$ and

$$h(z) = h(0) \frac{1}{1 - \xi z},$$

$$\tilde{g}(z) = (h(0)\,(-bz + d))^{-1} = \left(h(0)\, d\left(1 - \frac{b}{d} z\right)\right)^{-1}$$

so that

$$g = \frac{1}{h(0)^* d^*}\, e\left(\frac{b}{d}\right).$$

Thus

$$(Tf)(z) = (f, g)\, h(z) = \frac{1}{h(0)\, d}\left(f, e\left(\frac{b}{d}\right)\right) h(z) =$$

$$= \frac{1}{h(0)\, d} f\left(\frac{b}{d}\right) h(0) \frac{1}{1 - \xi z} = \frac{1}{cz + d} f\left(\frac{b}{d}\right).$$

Now suppose $A \in \mathscr{M}_0$, $T = m(A)$ and let A be singular. Since $d \neq 0$, we set $\xi = -c/d$ and obtain

$$(T\, p(x), p(y)) = (-bx + d)^{-1}\, (1 - \xi y^*)^{-1};$$

in particular, $|\xi| < 1$.

This makes it possible to compute $T\, p(x)$; indeed,

$$(T\, p(x), e(z)) = (T\, p(x), p(z^*)) =$$

$$= (-bx + d)^{-1}\, (1 - \xi z)^{-1} = (-bx + d)^{-1}\, (p(\xi), e(z))$$

so that

$$T\, p(x) = \frac{1}{-bx + d}\, p\left(-\frac{c}{d}\right).$$

To compute Te_k, consider a fixed $z \in D$ and observe that the numbers $(Te_k, e(z))$ satisfy, for every $x \in D$, the relation

$$\Sigma x^k (Te_k, e(z)) = (T\, p(x), e(z)) = \frac{1}{-bx + d} \frac{1}{1 + cz/d}.$$

For a fixed $z \in D$ this expression must be, as a function of x, an element of H^2.

It follows that $|b| < |d|$ and

$$(T p(x), e(z)) = \frac{1}{cz + d} \sum \left(\frac{b}{d} x\right)^k$$

so that

$$Te_k = \left(\frac{b}{d}\right)^k \frac{1}{cz + d}.$$

Given $f \in H^2$, we have thus

$$(Tf)(z) = \frac{1}{cz + d} f\left(\frac{b}{d}\right)$$

so that

$$T = he\left(\frac{b}{d}\right)^*$$

with $h(z) = (cz + d)^{-1}$.

Now suppose that $A \in \mathcal{M}_0$ and A is singular. Since $d \neq 0$ there exists a number ξ such that $c = \xi d$ and $a = \xi b$.

Thus $(ax - c)/(-bx + d) = -\xi$ for every $x \in D$.

Since $A \in \mathcal{M}_0$ we have, for $T = m(A)$,

$$T p(x) = \frac{1}{-bx + d} p(-\xi).$$

It follows that $T p(x)$ is multiple of the vector $p(-\xi)$ for every x. Since the linear combinations of the vectors $p(x)$, $x \in D$ are dense in H^2, the operator T is one-dimensional.

Now suppose that A satisfies the conditions of (3,3) and that A is nonsingular. We know that

$$\left|\frac{ax + b}{cx + d}\right| \leqq 1 \quad \text{for} \quad |x| < 1\,;$$

the inequality has to be strict, however, since otherwise the function

$$z \to \frac{az + b}{cz + d}$$

would assume its maximum in the interior of the unit disc and would, accordingly, be constant which is a contradiction with the assumption of nonsingularity of A.

Now we shall use a simple consequence of the closed graph theorem. If m is a mapping of D into itself such that $f \circ m$ belongs to H^2 for all $f \in H^2$ then the mapping

$$f \to f \circ m$$

is a bounded linear operator on H^2. Consider now a nonsingular matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

satisfying the conditions of (2,3). It follows from what has been said above that the mapping T which assigns to each $f \in H^2$ the function

$$\frac{1}{cz + d} f\left(\frac{az + b}{cz + d}\right)$$

is a bounded linear operator on H^2; it may also be characterized by the equality

$$T p(x) = \frac{1}{-bx + d} p\left(\frac{ax - c}{-bx + d}\right).$$

This reasoning proves the implications $3° \to 4° \to 1°$. The proof of the implication $1° \to 2°$ is simple. To complete the proof it suffices to observe that singularity of A implies that $m(A)$ is one-dimensional; this was proved in the first part of the present proposition. We have thus the remaining implication $2° \to 3°$ and the proof is complete.

In the rest of this section we intend to collect some simple propositions which describe the connection between the properties of a matrix $A \in \mathcal{M}_0$ and those of the corresponding operator $m(A)$.

(3,5) *If $A \in \mathcal{M}_0$ then $QA^*Q \in \mathcal{M}_0$ as well and*

$$m(A)^* = m(QA^*Q).$$

Proof. Given $x, y \in D$ we have

$$(m(A)^* p(y), p(x)) = (m(A) p(x), p(y))^* =$$
$$= ((-bx + d - y^*(ax - c))^{-1})^* = (c^*y + d^* - x^*(a^*y + b^*))^{-1} =$$
$$= \left(m\begin{pmatrix} a^* & -c^* \\ -b^* & d^* \end{pmatrix} p(y), p(x)\right) = (m(QA^*Q) p(y), p(x)).$$

(3,6) *For every $A \in \mathcal{M}_0$ we have $QA^{\mathsf{T}}Q \in \mathcal{M}_0$ and*

$$m(A)^{\mathsf{T}} = m(QA^{\mathsf{T}}Q).$$

Proof. According to lemma (1,1) we have

$$(m(A)^{\mathsf{T}} p(u), p(v)) = (m(A) p(v^*), p(u^*)) = (-bv^* + d - u(av^* - c))^{-1} =$$
$$= (cu + d - v^*(au + b))^{-1} = (m(QA^{\mathsf{T}}Q) p(u), p(v)).$$

(3,7) *For every $A \in \mathcal{M}_0$,*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

and $x \in D$

$$m(A)\, p(x) = \frac{1}{-bx+d}\, p\left(\frac{ax-c}{-bx+d}\right).$$

Proof. We have

$$(m(A)\, p(x), e(z)) = (m(A)\, p(x), p(z^*)) =$$

$$= (-bx+d-z(ax-c))^{-1} = \frac{1}{-bx+d}\left(1 - z\frac{ax-c}{-bx+d}\right)^{-1} =$$

$$= \sum_0^\infty \frac{(ax-c)^k}{(-bx+d)^{k+1}} z^k$$

so that

$$m(A)\, p(x) = \sum_0^\infty \frac{(ax-c)^k}{(-bx+d)^{k+1}} e_k.$$

(3,8) *The set* $\mathcal{M}_0$ *is closed with respect to matrix multiplication. If* A, B, P *are elements of* $\mathcal{M}_0$ *and* u, v *are given numbers in* D *then*

1° $AB \in \mathcal{M}_0$ *and* $m(B)\, m(A) = m(A\,B)$,

2° $AQP^*Q \in \mathcal{M}_0$ *and* $(m(A)\, p(v), m(P)\, p(u)) = (m(AQP^*Q)\, p(v), p(u))$.

Proof. We begin by proving the second assertion. Suppose that

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad P = \begin{pmatrix} p & q \\ r & s \end{pmatrix}.$$

We have then

$$(m(A)\, p(v),\ m(P)\, p(u)) =$$

$$= \left(\frac{1}{-bv+d}\, p\left(\frac{av-c}{-bv+d}\right),\ \frac{1}{-qu+s}\, p\left(\frac{pu-r}{-qu+s}\right)\right) =$$

$$= \frac{1}{(-bv+d)(-qu+s)^*} \frac{1}{1 - \dfrac{av-c}{-bv+d}\left(\dfrac{pu-r}{-qu+s}\right)^*} =$$

$$= ((-bv+d)(-qu+s)^* - (av-c)(pu-r)^*)^{-1} =$$

$$= ((-bs^* + ar^*)\, v + (ds^* - cr^*) - u^*((-bq^* + ap^*)\, v +$$

$$+ (dq^* - cp^*)))^{-1} = (m(W)\, p(v), p(u))$$

where

$$W = \begin{pmatrix} -bq^* + ap^* & bs^* - ar^* \\ -dq^* + cp^* & ds^* - cr^* \end{pmatrix}.$$

Clearly $W = AQP^*Q$.

The first assertion is an immediate consequence of the second one. We have

$$(m(A)\, m(B)\, p(v), p(u)) = (m(B)\, p(v), m(A)^*\, p(u)) =$$
$$= (m(B)\, p(v), m(QA^*Q)\, p(u)) =$$
$$= (m(BQ(QA^*Q)^*\, Q)\, p(v), p(u)) = (m(BA)\, p(v), p(u)) .$$

(3,9) *Suppose that $A \in \mathcal{M}_0$ and that the inverse $m(A)^{-1}$ exists. Then $|b| \leqq |a|$, $A^{-1} \in \mathcal{M}_0$ and*

$$m(A)^{-1} = m(A^{-1}) .$$

Proof. Suppose that $|b| > |a|$ so that the function g defined by $g(z) = (az + b)^{-1}$ belongs to H^2. Set $f = m(A)^{-1}\, g$; it follows that

$$\frac{1}{cz + d} f\left(\frac{az + b}{cz + d}\right) = \frac{1}{az + b}$$

for all $z \in D$ so that $f(y) = y^{-1}$ for every y which is of the form $(az + b)/(cz + d)$ for some $z \in D$. Since $m(A)$ is invertible, the matrix A is nonsingular so that such y fill some disc. This contradiction proves that $|b| \leqq |a|$.

Given $y \in D$, denote by f_y the element of H^2 for which

$$m(A) f_y = p(y) .$$

Thus

$$\frac{1}{cx + d} f_y \left(\frac{ax + b}{cx + d}\right) = \frac{1}{1 - xy}$$

for all $x \in D$. If z is of the form

$$z = \frac{ax + b}{cx + d}$$

then

$$x = \frac{dz - b}{a - cz}$$

and

$$f_y(z) = \frac{ad - bc}{a - cz - y(-b + dz)} = \frac{ad - bc}{a + by - z(c + dy)} .$$

Since $|a| \geqq |b|$ we have $a + by \neq 0$ and

$$f_y = \frac{ad - bc}{a + by}\, p\left(\frac{c + dy}{a + by}\right)$$

we see thus that

$$m(A)^{-1} = m(B)$$

where

$$B = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

so that $B = A^{-1}$.

Before stating the next lemma an explanation might be in order. If $a \in l^2$ we shall denote by $(p(\cdot), a)$ the function in H^2 whose value at the point $z \in D$ is

$$(p(z), a)\,;$$

clearly another way of describing this element of H^2 is

$$(p(\cdot), a) = \Sigma a_k^* e_k\,.$$

With this explanation in mind, we are ready to prove the following technical lemma

(3,10) *If a and b are given elements of l^2 then the scalar product of the H^2 functions $(p(\cdot), a)$ and $(p(\cdot), b)$ equals (b, a)*

$$((p(\cdot), a),\ (p(\cdot), b)) = (b, a)\,.$$

If A and B are operator matrices then

$$((A\ p(\cdot),\ p(u)),\ (B\ p(\cdot),\ p(v))) = (AB^*\ p(v),\ p(u))\,.$$

Proof. We have

$$((p(\cdot), a), (p(\cdot), b)) = (\Sigma a_k^* e_k,\ \Sigma b_j^* e_j) = \Sigma b_j a_j^* = (b, a)\,.$$

The second assertion is an immediate consequence of the first one for the case

$$a = A^*\ p(u)\,, \quad b = B^*\ p(v)\,.$$

The technical lemma just proved will be used to establish the following neat formula for Gram matrices of certain sequences of vectors in H^2.

(3,11) *Let $b_0, b_1, \ldots$ be a sequence in a Hilbert space H such that the mapping assigning to each $x \in H$ the sequence*

$$Tx = \{(x, b_j)\}_j$$

maps H into l^2. Then T is a bounded linear map of H into l^2 and the conjugate mapping T^* is characterized by the relation

$$T^* e_k = b_k\,.$$

The matrix of the mapping TT^* with respect to the natural basis of l^2 is the Gram matrix of the sequence b_j.

Proof. For each $x \in H$ and each $k = 0, 1, \ldots$

$$(x,\ T^* e_k) = (Tx,\ e_k) = (x,\ b_k)$$

so that $T^*e_k = b_k$. Furthermore, given i, k, we have

$$\mathcal{M}(TT^*)_{ik} = (TT^*e_k, e_i) = (T^*e_k, T^*e_i) = (b_k, b_i) = G(b_0, b_1, \ldots)_{ik}\,.$$

(3,12) *The bilinear form corresponding to* T^* *is*

$$(T^*\, p(x), p(y)) = \Sigma x^k\, b_k(y^*)\,.$$

Proof. $(T^*\, p(x), p(y)) = (T^*\, \Sigma x^k e_k, e(y^*)) = \Sigma x^k (T^* e_k, e(y^*)) = \Sigma x^k\, b_k(y^*)$.

(3,13) *Suppose we have a sequence* $b_j \in H^2$ *of the form*

$$b_j(z) = \frac{(az + b)^j}{(cz + d)^{j+1}}$$

such that the matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

belongs to $\mathcal{M}_0$. *Then*

$$G(b_0, b_1, .\ ) = m(AQA^*Q)\,.$$

Proof. Denote by R the operator defined by $Re_k = b_k$. Then

$$(R\, p(x), p(y)) = (\Sigma x^k b_k, p(y)) = \Sigma x^k\, b_k(y^*) =$$

$$= \frac{1}{cy^* + d}\left(1 - x\,\frac{ay^* + b}{cy^* + d}\right)^{-1} = (cy^* + d - x(ay^* + b))^{-1}$$

so that $R = m(A)$. According to (3,11)

$$G(b_0, b_1, \ldots) = m(R^*R) = m(A)^*\, m(A) = m(QA^*Q)\, m(A) = m(AQA^*Q)\,.$$

4. ORTHONORMALIZATION

We intend to show in this section how the technical machinery described in the preceding paragraph may be used to express the operator

$$S \mid \operatorname{Ker}(S - \alpha)^n \quad (|\alpha| < 1)$$

as a matrix with respect to an orthonormal basis.

To find this matrix we consider first a natural basis of the space $\operatorname{Ker}(S - \alpha)^n$ with respect to which the restriction of S has a simple matrix; the result of the preceding paragraph may then be used to write down the matrix of this operator with respect to the basis obtained by the Gram-Schmidt process from the given one.

Suppose we are given a sequence $b_0, b_1, \ldots$ of linearly independent elements of H^2 and an operator $A \in B(H^2)$. We shall suppose that the closed linear span E

of the sequence is invariant with respect to A. Denote by $f_0, f_1, f_2, \ldots$ the orthonormal sequence obtained from $b_0, b_1, b_2, \ldots$ by the Gram-Schmidt process. If $f_s = \Sigma V_{ts} b_t$ then the matrix of $A \mid E$ with respect to the basis $f_0, f_1, f_2, \ldots$ is

$$V^{-1}KV$$

if K stands for the matrix for which

$$Ab_j = \Sigma K_{tj} b_t .$$

We shall restrict ourselves to the case where K is column finite – a hypothesis which is fulfilled in the concrete case to be investigated in this section. The matrix V is upper triangular by its very nature so that V^{-1} possesses the same property; accordingly, no convergence problems arise in the matrix product $V^{-1}KV$. Since we shall deal with column-finite matrices K only there will be no danger of misunderstanding if we extend to this situation the terminology used in the finite-dimensional case: we shall call K the "matrix of A with respect to the basis b_j"; the word basis used here does not imply, of course, that any additional properties of the sequence b_j are assumed except that the b_j are linearly independent.

Now consider the particular case of the operator S and the linearly independent sequence

$$b_j(z) = \frac{z^j}{(1 - \alpha z)^{j+1}}, \quad j = 0, 1, 2, \ldots .$$

The matrix of S with respect to the basis $b_0, b_1, \ldots$ is fairly simple since

$$Sb_k = \alpha b_k + b_{k-1}$$

for $k = 0, 1, 2, \ldots$ if we agree to set $b_{-1} = 0$. Hence the matrix of S equals $\alpha I + N$ where N is the matrix

$$N = \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots \\ 0 & 0 & 1 & 0 & \ldots \\ 0 & 0 & 0 & 1 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix}.$$

Our task reduces thus to finding the matrix of S with respect to the basis f_j. We observe first that the same sequence f_j is obtained if the sequence b_j to be orthonormalized is replaced by a sequence of the form $w_j b_j$ with nonzero w_j. We shall choose $w_k = \beta^k$ for a suitable β – in this manner we obtain a sequence $g_j = \beta^j b_j$ to which the foregoing theory may be applied; the choice $\beta = 1 - |\alpha|$ will do.

Upon multiplying the above relation by β^k we obtain

$$Sg_k = \alpha g_k + \beta g_{k-1} .$$

If we denote by W the transition matrix then

$$f_j = \sum w_{rj} g_r ;$$

since each f_j is a linear combination of $b_0, \ldots, b_j$ we shall have $w_{rj} = 0$ for $r > j$ so that W will be upper triangular and will satisfy

$$1 = W^* G(g_0, g_1, \ldots) W .$$

Thus the matrix of S with respect to the g_j will be

$$\alpha I + \beta N .$$

To construct such a matrix W, it will be sufficient to find an upper triangular B for which $G = B^* B$ and then set $W = B^{-1}$.

We intend to look for a B of the form $m(A)$ for a suitable $A \in \mathcal{M}_0$.

Now $m(A)$ should be upper triangular, in other words, $m(A)_{jk}$ should be zero for $j > k$ so that the corresponding bilinear form should not contain products $x^k y^{*j}$ for $k < j$; since y^{*j} is multiplied by $(ax - c)^j/(-bx + d)^{j+1}$ this is only possible if $c = 0$. We are thus looking for a matrix M of the form

$$M = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

such that $m(M)^* \, m(M) = G(g_0, g_1, \ldots)$.

According to the remarks following (3,11)

$$G(g_0, g_1, \ldots) = m(AQA^*Q)$$

for

$$A = \begin{pmatrix} \beta & 0 \\ -\alpha & 1 \end{pmatrix}$$

so that

$$G(g_0, g_1, \ldots) = m \begin{pmatrix} |\beta|^2 & \beta\alpha^* \\ -\alpha\beta^* & 1 - |\alpha|^2 \end{pmatrix} .$$

The requirement on M is thus

$$\begin{pmatrix} |a|^2 - |b|^2 & bd^* \\ -db^* & |d|^2 \end{pmatrix} = \begin{pmatrix} |\beta|^2 & \beta\alpha^* \\ -\alpha\beta^* & 1 - |\alpha|^2 \end{pmatrix} .$$

Denoting by q the difference $1 - |\alpha|^2$ and by $q^{1/2}$ the positive square root of q, we obtain

$$M = \begin{pmatrix} \varepsilon_2 \beta q^{-1/2} & \varepsilon_1 \beta \alpha^* q^{-1/2} \\ 0 & \varepsilon_1 q^{1/2} \end{pmatrix}$$

$$M^{-1} = \begin{pmatrix} \varepsilon_2^* \beta^{-1} q^{1/2} & -\varepsilon_2^* \alpha^* q^{-1/2} \\ 0 & \varepsilon_1^* q^{-1/2} \end{pmatrix}$$

for suitable numbers $\varepsilon_1, \varepsilon_2$ of modulus one.

To compute the matrix

$$m(M)\,N\,m(M)^{-1}$$

we consider the corresponding bilinear form

$$(m(M)\,N\,m(M^{-1})\,p(x),\,p(u)) = (S\,m(M^{-1})\,p(x),\,m(M)^*\,p(u))\,.$$

We shall use the formula obtained in (3,4). If $T \in \mathscr{M}_0$ and

$$T = \begin{pmatrix} r & s \\ t & w \end{pmatrix}$$

then

$$m(T)\,p(x) = \frac{1}{-sx + w}\,p\left(\frac{rx - t}{-sx + w}\right);$$

at the same time $S\,p(a) = a\,p(a)$ for every $a \in D$. Thus

$$S\,m(T)\,p(x) = \frac{1}{-sx + w}\,\frac{rx - t}{-sx + w}\,p\left(\frac{rx - t}{-sx + w}\right) =$$

$$= \frac{rx - t}{-sx + w}\,m(T)\,p(x)\,.$$

Using these facts, we obtain

$$(m(M)\,N\,m(M)^{-1}\,p(x),\,p(u)) = H(m(M^{-1})\,p(x),\,m(M)^*\,p(u)) =$$

$$= H(p(x),\,p(u)) = H\,\frac{1}{1 - u^*x}$$

where

$$H = \frac{1}{\beta}\,\frac{\varepsilon_2^* q^{1/2} x}{\varepsilon_2^* \alpha^* q^{-1/2} x + \varepsilon_1^* q^{-1/2}} = \frac{q}{\beta}\,\frac{x}{\varepsilon_1^* \varepsilon_2 + \alpha^* x} = -\frac{q}{\beta}\,\frac{1}{\omega^*}\,\frac{x}{1 - \omega\alpha^* x}$$

if we set $\omega = -\varepsilon_1\varepsilon_2^*$. The bilinear form corresponding to the matrix of $(S - \alpha)$ with respect to the orthonormal system $f_0, f_1, \ldots$ assumes thus the following simple form

$$-q\omega\,\frac{x}{1 - \omega\alpha^* x}\,\frac{1}{1 - u^* x} = -q\omega\,\Sigma(\omega\alpha^*)^k\,\frac{x^{k+1}}{1 - u^* x}\,.$$

Since multiplication by x amounts to the same as shifting the matrix one column to the right the corresponding matrix is the Toeplitz matrix T which has the number $-q\omega^k\alpha^{*k-1}$ on the k-th superdiagonal, $(k = 1, 2, \ldots)$, in other words

$$t_{rs} = -q\omega^{s-r}(\alpha^*)^{s-r-1}$$

for $s - r = 1, 2, \ldots,\ r \geqq 0$.

Now we are ready to state a result which summarizes our investigations.

(4,1) *Let α be a complex number with $|\alpha| < 1$. Consider the infinite Toeplitz matrix*

$$T(\alpha) = \begin{pmatrix} \alpha & -q & -q\alpha^* & -q\alpha^{*2} & -q\alpha^{*3} & \dots \\ 0 & \alpha & -q & -q\alpha^* & -q\alpha^{*2} & \dots \\ 0 & 0 & \alpha & -q & -q\alpha^* & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

and denote by $T(\alpha, n)$ the n by n matrix consisting of the $T(\alpha)_{j,k}$ for $0 \leqq j, k \leqq \leqq n - 1$.

Given any natural number n, there exists an orthonormal basis of the space $\operatorname{Ker}(S - \alpha)^n$ *with respect to which the operator* $S \mid \operatorname{Ker}(S - \alpha)^n$ *has matrix* $T(\alpha, n)$.

Combining this with the main result of [3] we obtain the following.

(4,2) Corollary. *Let n be a natural number and r a nonnegative number, $r < 1$. Consider the family of all linear operators A on an n-dimensional Hilbert space such that*

1° *the norm of A does not exceed* 1

2° *the spectral radius of A does not exceed r.*

Then the maximum of $|A^n|$ is attained for the operator $T(r, n)$.

References

[1] *V. Pták:* Norms and the spectral radius of matrices. Czechosl. Math. J. *87* (1962), 553—557.

[2] *V. Pták:* Rayon spectral, norme des itérés d'un opérateur et exposant critique. C.R. Acad. Sci. Paris 265 (1967), 257—269.

[3] *V. Pták:* Spectral radius, norms of iterates and the critical exponent. Lin. Alg. Appl. *1* (1968), 245—260.

[4] *V. Pták:* Isometric parts of operators and the critical exponent. Čas. pěst. mat. *101* (1976), 383—388.

[5] *V. Pták:* Universal estimates of the spectral radius. Proceedings of the semester on spectral theory, Banach Center Publ., vol. *8*, (Spectral theory) (1982), 373—387.

[6] *V. Pták:* A lower bound for the spectral radius. Proceedings of the American Mathematical Society *80* (1980), 435—440.

[7] *V. Pták:* A maximum problem for matrices. Lin. Alg. Appl. *28* (1979), 193—204.

[8] *V. Pták:* Functions of operators and the spectral radius († *N. J. Young*). Lin. Alg. Appl. *29* (1980), 357—392.

[9] *D. Sarason:* Generalized interpolation in H^∞. Trans. Amer. Math. Soc. *127* (1967), 179—203.

[10] *B. Sz.-Nagy:* Sur la norme de certains opérateurs. Acta Math. Acad. Sci. Hungar. *20* (1969), 331—334.

[11] *B. Sz.-Nagy, C. Foias:* Analyse harmonique des opérateurs. Akademiai Kiadó 1967.

[12] *N. J. Young:* Analytic programmes in matrix algebras. Proc. Lond. Math. Soc. *36* (1978), 226—242.

Author's address: 115 67 Praha 1, Žitná 25 (Matematický ústav ČSAV).