Biorthogonal Systems and the Infinite Companion Matrix

Vlastimil Pták
Institute of Mathematics
Czechoslovak Academy of Sciences
Žitná 25
115 67 Praha 1, Czechoslovakia

Submitted by Miroslav Fiedler

ABSTRACT

A nonsymmetric analogue of a Gram matrix is used to represent the infinite companion matrix introduced by the author in [4] as a projection operator in H^2. The method consists in constructing, for a given set $f_0,\dots,f_{n-1}$ of linearly independent vectors, another set $g_0,\dots,g_{n-1}$ which lies in a given space and satisfies $(f_i, g_k)=\delta_{ik}$. The infinite companion matrix may then be interpreted as an interpolation operator. Connections with Lyapunov-type equations are also explained.

INTRODUCTION

Consider a contraction T on a Hilbert space H. If the spectral radius of T is less than one, then $T^r \to 0$ as r tends to infinity, so that there will be a power of T whose norm is less than one. In 1960 the present author discovered the following rather surprising fact: if the Hilbert space has dimension n, then already $|T^n|$ is smaller than one [2]. It is somewhat more difficult to give this result a quantitative form: to find the supremum of $|T^n|$ if T ranges over all contractions T whose spectral radius does not exceed a given number $p<1$. The solution was given in the second paper of the series [3, 4]—a description was given of a contraction T with spectral radius r for which $|T^n|$ assumes its maximum. The solution of this maximum problem was based on the consideration of individual sequences $x_0, Tx_0, T^2x_0,\dots$. Since T is annihilated by its characteristic polynomial, such sequences are determined by the first n terms $x_0, Tx_0,\dots,T^{n-1}x_0$. For a contraction the Gram matrix G of these vectors satisfies an equation of Lyapunov type

$$G - C^*GC = P$$

for a certain positive semidefinite P. The maximum problem was then transformed into an extremum problem for the matrices P.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/83/010057-22$3.00

In spite of the fact that subsequent investigations (notably the paper of B. Sz.-Nagy [7]) have yielded a more elegant proof, the original idea has not lost its interest, in particular because it puts into evidence the connection between the maximum problem and Lyapunov-type equations.

In order to describe the relation between the infinite sequence $x_0, Tx_0, \ldots$ and its first n initial conditions, the author introduced, in [4], the so called "infinite companion matrix," whose study forms the subject matter of the present note. We intend to elaborate more deeply the ideas used in the author's original proof, in particular the application of Gram matrices. Of course, Gram matrices can only be used to represent positive semidefinite matrices, while some of the matrices naturally occurring in the theory are even nonsymmetric. Also, subsequent investigations of N. J. Young [8] have shown that nonperpendicular projections can be used with advantage in the theory. This led to the consideration of generalized Gram matrices which make it possible to present a geometrical interpretation of the results.

We obtain a unified treatment of several results studied separately thus far, as well as a considerable simplification of the proofs. Among other results, we give a quite simple and geometrically intuitive proof of an explicit formula for the solution of Lyapunov equations due to N. J. Young [8]. The use of dual n-dimensional subspaces reveals also the geometrical meaning of the function F used by the author [5] in order to obtain an explicit expression for the infinite companion matrix. Also, the function F may be used in a simple manner to obtain the classical interpolation formula.

1. GENERALIZED GRAM MATRICES

Consider a Hilbert space H. If a vector $u \in H$ is represented as a linear combination of n given vectors $f_0, \ldots, f_{n-1}$,

$$u = x_0 f_0 + \cdots + x_{n-1} f_{n-1},$$

then its scalar product with the vector

$$v = y_0 g_0 + \cdots + y_{n-1} g_{n-1}$$

equals

$$(u, v) = \left(\sum_k x_k f_k, \sum_j y_j g_j \right)$$
$$= \sum_{k,j} y_j^* (f_k, g_j) x_k = \sum y_j^* g_{jk} x_k.$$

We have denoted by g_{jk} the scalar product (f_k, g_j). The matrix with elements g_{jk} will be denoted by $G(f, g)$ and will be called the Gram matrix of the n-tuples $f = (f_0, \dots, f_{n-1})$ and $g = (g_0, \dots, g_{n-1})$. If the coordinates are interpreted as column vectors in C^n,

$$x = (x_0, \dots, x_{n-1})^T, \qquad y = (y_0, \dots, y_{n-1})^T,$$

the above relation may be rewritten in the form

$$(u, v) = y^* G x = (Gx, y).$$

We shall represent an n-tuple of vectors as a row vector $f = (f_0, \dots, f_{n-1})$. If $g = (g_0, \dots, g_{n-1})$ is another such n-tuple, we can write formally

$$G(f, g) = g^* f.$$

In the particular case of vectors in C^n, the row vector b may be identified with the n-by-n matrix $(f_0, \dots, f_{n-1})$, and it is not difficult to verify that the above formula remains true even in this interpretation of $g^* f$.

Consider now two coordinate vectors

$$x = (x_0, \dots, x_{n-1})^T, \qquad y = (y_0, \dots, y_{n-1})^T$$

and the vectors

$$u = \sum x_j f_j, \qquad v = \sum y_j g_j.$$

Writing them in the form $u = fx$, $v = gy$, their scalar product becomes $(u, v) = (fx, gy) = (gy)^* fx = y^* g^* fx = y^* G x = (Gx, y)$.

Suppose we have two operators $A, B \in \alpha(H)$ such that the matrices $M(A)$ with elements a_{ik} and $M(B)$ with elements b_{ik} satisfy

$$Af_i = \sum_s a_{si} f_s,$$

$$Bg_j = \sum b_{rj} g_r.$$

Then $G(Af, Bg) = M(B)^* G(f, g) M(A)$.

Suppose we have two n-tuples $f_0, \ldots$ and $g_0, \ldots$ such that $G(f, g)$ is invertible. The following simple method of constructing the inverse matrix $G(f, g)^{-1}$ will be used in the sequel. Suppose we find two operators A and B such that

$$(Af_i, Bg_k) = \delta_{ik};$$

let us show that $G^{-1} = M(A)M(B)^*$. Indeed, it follows from the above formula that

$$I = M(B)^* G M(A),$$

whence

$$M(B)^*[GM(A)M(B)^* - 1] = [M(B)^* GM(A)]M(B)^* - M(B)^* = 0.$$

In a similar manner

$$[M(A)M(B)^* G - 1]M(A) = M(A)[M(B)^* GM(A) - 1] = 0.$$

Let F, G be two finite-dimensional subspaces of H. We shall say that F and G are dual to each other if, for each $f \in F$, $f \neq 0$, there exists a $g \in G$ such that $(f, g) \neq 0$ and for each $g \in G$, $g \neq 0$ there exists an $f \in F$ with $(f, g) \neq 0$. Clearly this is only possible if F and G have the same dimension.

The following equivalence is obvious:

PROPOSITION 1.1. *Let F and G be two finite-dimensional subspaces of H. Then these are equivalent*:

(1) *F and G are dual to each other*,
(2) *$G \cap F^{\perp} = 0$ and $F \cap G^{\perp} = 0$*,
(3) *for every basis $f = f_0, \ldots, f_{n-1}$ of F and every basis $g = g_0, \ldots, g_{n-1}$ of G, the matrix $G(f, g)$ is nonsingular*,
(4) *there exists a basis f for F and a basis g for G such that* $\det G(f, g) \neq 0$,
(5) *$F + G^{\perp} = H$ and $G + F^{\perp} = H$.*

Proof. Condition (2) is nothing more than a reformulation of (1).

The equivalence of (2), (3), and (4) is elementary linear algebra. Now suppose that (4) is satisfied, and consider the basis $g_0, \ldots, g_{n-1}$ of G and the basis $f_0, \ldots, f_{n-1}$ of F. Since $\det(g_i, f_j)$ is different from zero, it is possible to determine, for each $x \in H$, coefficients $\alpha_0, \ldots, \alpha_{n-1}$ such that

$$\left(x - \sum \alpha_j g_j, f_k\right) = 0$$

for $k = 0, 1, \ldots, n-1$. The vector $g = \Sigma \alpha_j g_j$ satisfies the conditions $g \in G$ and $x - g \in F^{\perp}$, so that $x = g + (x - g) \in G + F^{\perp}$. In a similar manner we see that $H = F + G^{\perp}$. On the other hand, assume (5) and consider an $x \in G \cap F^{\perp}$. Since $x = f + g^{\perp}$ for suitable $f \in F$ and $g^{\perp} \in G^{\perp}$, we have $(x, x) = (x, f) + (x, g^{\perp})$, and both summands are zero. Hence $x = 0$. ∎

Proposition 1.2. *Let F and G be a pair of dual subspaces of H, and let $f = f_0, \ldots, f_{n-1}$ be a fixed basis of F. For a set of vectors $h = h_0, \ldots, h_{n-1}$ the following conditions are equivalent:*

(1) *$x - \Sigma(x, h_j) f_j \in G^{\perp}$ for every $x \in H$,*
(2) *$h_j \in G$ and $G(f, h) = 1$,*
(3) *$h_j = \Sigma_s d_{sj} g_s$, where $g = (g_0, \ldots, g_{n-1})$ is a basis of G and d_{sj} are elements of the matrix $G(g, f)^{-1}$.*

For a fixed basis f there exists exactly one system h with the properties above. In particular,

$$f = \sum (f, h_j) f_j$$

for all $f \in F$.

Proof. Suppose first that (1) is satisfied, and let us show that the h_j lie in G. To this end choose a basis $g_0, g_1, \ldots, g_{n-1}$ of the space G. For each $x \in H$ we have $x - \Sigma(x, h_j) f_j \in G^{\perp}$. It follows that, for every $x \in H$ and every s,

$$\begin{aligned}\left(x, g_s - \sum (g_s, f_k) h_k\right) &= (x, g_s) - \sum (f_k, g_s)(x, h_k) \\ &= \left(x - \sum (x, h_k) f_k, g_s\right) = 0,\end{aligned}$$

so that each g_s is a linear combination of the h_k,

$$g_s = \sum (g_s, f_k) h_k,$$

and, the matrix $G(g, f)$ being nonsingular, the h_k are linear combinations of the g_s; hence $h_k \in G$.

For every $f \in F$ we have

$$f - \sum (f, h_j) f_j \in F \cap G^{\perp};$$

hence $f = \Sigma(f, h_j)f_j$. In particular $f_k = \Sigma_j(f_k, h_j)f_j$, whence $G(f, h) = 1$. Now write the h_j in the form $h_j = \Sigma_s d_{sj} g_s$. We have $1 = G(h, f)$ and $\delta_{jk} = (h_j, f_k) = \Sigma_s d_{sj}(g_s, f_k) = \Sigma_s d_{sj} G(g, f)_{ks}$, so that $G(g, f)D = 1$, where $D = (d_{sj})$. We have thus shown that there exists at most one system h with property (1). Any such system must be of the form $h_k = \Sigma d_{sk} g_s$ with $D = G(g, f)^{-1}$. To prove the existence, take the vectors $h_k = \Sigma_s d_{sk} g_s$ with $D = G(g, f)^{-1}$; we see first that they belong to G and that $G(f, h) = 1$. Also, $f = \Sigma(f, h_k)f_k$ for every $f \in F$. Let us show now that (1) holds for every $x \in H$. Given x, we may write $x = f + g^{\perp}$ with $f \in F$, $g^{\perp} \in G^{\perp}$, so that $(x, h_k) = (f, h_k)$ for all k. Furthermore

$$x - \sum (x, h_k) f_k = f + g^{\perp} - \sum (f, h_k) f_k = g^{\perp} \in G^{\perp}. \qquad \blacksquare$$

The proposition just proved shows that—given a dual pair F, G and a basis $f = f_0, \ldots, f_{n-1}$ of F—there exist among all systems $m = m_0, \ldots, m_{n-1}$ with the property

$$G(f, m) = I$$

exactly one for which all $m_j \in G$. This system, denoted by $h = h_0, \ldots, h_{n-1}$ has the property that the mapping $\Pi(F, G^{\perp})$ defined by

$$x \to \sum (x, h_j) f_j$$

is a projection with range F and kernel $G^{\perp}$.

It is not difficult to verify the following general formula:

$$\Pi(F, G^{\perp})^* = \Pi(G, F^{\perp}).$$

The fact that among all systems orthogonal to a given basis f of F exactly one may be singled out which is contained in G will be formulated as a lemma.

LEMMA 1.3. *Given a dual pair F, G and a basis $f = (f_0, \ldots, f_{n-1})$ of F, there exists exactly one system $h = (h_0, \ldots, h_{n-1})$ satisfying the conditions*

(1) $G(f, h) = I$,

(2) $h_j \in G$ *for all* j.

In our applications the spaces F and G will be of the form $F = \operatorname{Ker} A$ and $G = \operatorname{Ker} B$. In this case a basis-free expression for $\Pi(G, F^{\perp})$ may easily be given. It is based on a simple observation which we prefer to formulate in its full generality for infinite-dimensional Hilbert spaces. The following two lemmata might serve as an explanation of its algebraic substance.

LEMMA 1.4. *Consider three linear spaces H, K, M and two linear operators*

$$B: H \to M,$$

$$C: K \to H.$$

Suppose there exists a linear operator $Y: M \to K$ such that YBC is the identity operator on K. Set

$$P = CYB, \qquad Q = BCY.$$

Then

P is the projection operator in H onto $\operatorname{Im} C$ *along* $\operatorname{Ker} YB$,
Q is the projection operator in M onto $\operatorname{Im} BC$ *along* $\operatorname{Ker} Y$.

Proof. It is easy to verify that $P^2 = P$ and $Q^2 = Q$. Clearly $\operatorname{Im} P \subset \operatorname{Im} C$; the inclusion $\operatorname{Im} C \subset \operatorname{Im} P$ is a consequence of the identity $PC = C$. Since $YBP = YB$, we have $YB(I - P) = 0$, so that $\operatorname{Ker} P \subset \operatorname{Ker} YB$. The inclusion $\operatorname{Ker} YB \subset \operatorname{Ker} P$ is obvious. The assertion about Q follows in a similar manner from the identities $YQ = Y$ and $QBC = BC$. ∎

LEMMA 1.5. *Suppose that H, K, M are three Hilbert spaces, and consider two linear operators*

$$A: H \to K, \qquad B: H \to M.$$

Suppose there exists a linear operator Y such that YBA^* is the identity operator on K. Suppose that $H = \operatorname{Im} A^* + \operatorname{Ker} B$. *Then A^*YB is the projection on H onto* $\operatorname{Im} A^*$ *along* $\operatorname{Ker} B$.

Proof. According to the preceding result, A^*YB is the projection of H onto $\operatorname{Im} A^*$ along $\operatorname{Ker} YB$. Hence our assertion will be proved if we show that $\operatorname{Ker} YB = \operatorname{Ker} B$. To prove the inclusion $\operatorname{Ker} YB \subset \operatorname{Ker} B$, consider an x for which $YBx = 0$. Now x may be written in the form $x = A^*u + z$ for some u and some $z \in \operatorname{Ker} B$; it follows that

$$u = YBA^*u = YB(x - z) = 0$$

so that $x = z \in \operatorname{Ker} B$. The inclusion $\operatorname{Ker} B \subset \operatorname{Ker} YB$ being immediate, this proves the proposition. ■

The result which we shall need is little more than a modification of the preceding lemma.

PROPOSITION 1.6. *Suppose A and B are two bounded linear operators on the Hilbert space H, such that* $\operatorname{Im} A^* + \operatorname{Ker} B$ *is dense in H. Suppose that there exists a bounded linear operator Y such that $YBA^* = I$. Then* $\operatorname{Im} A^*$ *is closed, and*

$$P = A^*YB$$

is the projection $\Pi(\operatorname{Ker} A^{\perp}, \operatorname{Ker} B)$ *onto* $\operatorname{Im} A^*$ *along* $\operatorname{Ker} B$.

Proof. The estimate

$$|x| = |YBA^*x| \leqslant |YB||A^*x|$$

implies that the range of A^* is closed.

We have $P^2 = A^*YBA^*YB = A^*(YBA^*)YB = P$, so that P is a projection. Since $PA^* = (A^*YB)A^* = A^*(YBA^*) = A^*$, the operator P leaves invariant elements of the range of A^*.

Clearly $\operatorname{Ker} B \subset \operatorname{Ker} P$. Let us show that the two kernels are identical. We show first that $\operatorname{Ker} P \subset \operatorname{Ker} YB$. Indeed, $YB = (YBA^*)YB = YB(A^*YB) = YBP$, so that $Px = 0$ implies $YBx = 0$.

Now we use the density of $\operatorname{Im} A^* + \operatorname{Ker} B$ to show that $\operatorname{Ker} YB \subset \operatorname{Ker} B$. Suppose that $YBx = 0$, and let $\varepsilon > 0$ be given. Then there exist elements z and k such that

$$|x - A^*z - k| < \varepsilon$$

and $Bk = 0$. Hence

$$\begin{aligned} z = YBA^*z &= YB(x - k) - YB(x - A^*z - k) \\ &= -YB(x - A^*z - k), \end{aligned}$$

so that $|z| \leqslant |YB|\varepsilon$. It follows that

$$\begin{aligned}|Bx| &\leqslant |B(A^*z+k)|+|B(x-A^*z-k)|\\ &\leqslant |BA^*z|+|B|\varepsilon \leqslant |BA^*||z|+|B|\varepsilon\\ &\leqslant |BA^*||YB|\varepsilon+|B|\varepsilon,\end{aligned}$$

so that $|Bx|$ may be made arbitrarily small. Hence $Bx=0$, and the proof is complete. ∎

We conclude this section by pointing out how these facts may be used to obtain the inverse of a matrix $G(f,g)$. First of all, we express the inverse of $G(f,g)$ in the form $G(b,a)$ for suitable systems a,b and then indicate a simple method of practical construction of such systems.

Proposition 1.7. *Let F,G be a dual pair with bases $f=(f_0,\ldots,f_{n-1})$ and $g=(g_0,\ldots,g_{n-1})$. Suppose we are given two systems*

$$a=(a_0,\ldots,a_{n-1}),$$
$$b=(b_0,\ldots,b_{n-1}),$$

orthogonal respectively to f and g, so that $G(f,a)=I$, $G(g,b)=I$. If $b_j\in F$ for all j, then

$$G(f,g)G(b,a)=I.$$

Proof.

$$\begin{aligned}(b_k,g_i)&=\Big(\sum_r (b_k,a_r)f_r,g_i\Big)\\ &=\sum_r (f_r,g_i)(b_k,a_r)=\sum_r G(f,g)_{ir}G(b,a)_{rk}.\end{aligned}$$ ∎

Given a dual pair F,G and a basis $f_0,\ldots,f_{n-1}$ of F, there exists, as we have seen, exactly one system $h_0,\ldots,h_{n-1}$ of vectors in G for which $G(f,h)=I$. It may be constructed from any basis $g_0,\ldots,g_{n-1}$ of G by setting

$$h_j=\sum w_{sj}g_s,$$

the w_{sj} being elements of the matrix $W = G(g, f)^{-1}$. This is not satisfactory from the practical point of view. In many considerations the following construction turns out to be useful. Take any set of vectors u_j such that $G(f, u) = 1$, and set $h_j = \Pi(G, F^{\perp})u_j$. We have thus $u_j = h_j + k_j$, where $h_j \in G$ and $k_j \in F^{\perp}$, whence $(f_k, h_j) = (f_k, u_j)$. It follows that

$$G(f, h) = G(f, u) = I.$$

2. THE INFINITE COMPANION MATRIX

In what follows the space H will be the Hardy space H^2 of holomorphic functions on the unit disc. We shall denote by $e_0, e_1, \ldots$ the orthonormal system of functions $e_n(z) = z^n$. For each y, $|y| < 1$, let $e(y)$ be the element of H^2 given by

$$e(y)(z) = \frac{1}{1 - y^* z}.$$

We have thus $(f, e(y)) = f(y)$ for each $f \in H^2$. If $x \in H^2$, we denote by $\tilde{x}$ the function

$$\tilde{x}(z) = [x(z^*)]^*.$$

If p is a polynomial of degree n, we define p_0 and p_1 as

$$p_0(z) = z^n \left[p\left(\frac{1}{z^*} \right) \right]^*, \qquad p_1(z) = z^n p\left(\frac{1}{z} \right),$$

so that

$$p_0 = (\tilde{p})_1, \qquad p_1 = (p_0)^{\sim} = (\tilde{p})_0.$$

The (backward) shift operator S is defined as the mapping which assigns to a function f the function g as follows:

$$g(z) = \frac{1}{z}[f(z) - f(0)].$$

Its adjoint S^* is the operator of multiplication by z.

In the course of his investigations of the connection between the norms of powers of an operator and its spectral radius, the present author introduced [4], for each polynomial p, an infinite companion matrix. Let us recapitulate briefly its definition and properties. It was originally denoted by T^{∞}. Since we shall consider, occasionally, more than one polynomial, we shall write $T^{\infty}(p)$ to mark the dependence on p. Also, it is convenient to change the numbering of the indices slightly. Given a polynomial p of degree n, written

$$p(z) = -\left(a_0 + \cdots + a_{n-1}z^{n-1}\right) + z^n,$$

the companion matrix $C(p)$ of p is

$$C(p) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1 \\ a_0 & a_1 & a_2 & \cdots & a_{n-1} \end{pmatrix}.$$

The matrix $T^{\infty}(p)$ has n columns numbered $0,1,\ldots,n-1$ and an infinite number of rows $0,1,2,\ldots,$ with the following properties (each of which is characteristic):

(1) the jth column c_j of $T^{\infty}(p)$ is the solution of the recurrence relation

$$x_{r+n} = a_0x_r + a_1x_{r+1} + \cdots + a_{n-1}x_{r+n-1}$$

with the initial conditions

$$x_j = 1$$

and

$$x_k = 0 \qquad \text{for all } 0 \leqslant k \leqslant n-1 \text{ different from } j,$$

(2) given any $r = 0,1,2,\ldots,$ the matrix consisting of the n consecutive rows of $T^{\infty}(p)$ starting with the rth row is equal to $C(p)^r$,

(3) given any n-dimensional operator A such that $p(A) = 0$ and any nonnegative r, then

$$A^r = \sum_{k=0}^{n-1} t_{rk}A^k,$$

(4) if $F(y, z)$ stands for the function $\sum_{r=0}^{\infty}\sum_{j=0}^{u-1} t_{rj} y^j z^r$, then

$$F(y,z)=\frac{1}{1-yz}\left(1-\frac{p(z)}{p(1/y)}\right)$$

(this formula makes it possible to express the t_{rj} in terms of the coefficients a_k),

(5) if $\alpha_1,\dots,\alpha_n$ are the roots of p, then

$$t_{rj}=(-1)^{n-j-1}\sum\binom{q(e_1,\dots,e_n)-1}{n-j-1}\alpha_1^{e_1}\cdots\alpha_n^{e_n},$$

the summation ranging over all n-tuples of nonnegative integers $e_1,\dots,e_n$ such that $\sum e = r-j$, while $q(e_1,\dots,e_n)$ stands for the number of those e which are positive.

The formula in (5) is due to Z. Dostál [1] and V. Pták [5]. The function $F(y, z)$ appears first in [5]; some of its further properties will be discussed below.

Let us remark here that the results of the present note make it possible to give yet another interpretation to the entries of T^{∞}:

(6) The rth row of T^{∞} consists of the coefficients of the polynomial obtained as remainder upon dividing x^r by $p(x)$.

Consider now a fixed polynomial p:

$$p(z)=a_0+a_1z+\cdots+a_nz^n=a_n(z-\alpha_1)\cdots(z-\alpha_n),$$

with all roots inside the unit disc. Every element h of the space H^2 may be written in the form

$$h=h_1+h_2,$$

where h_1 is a polynomial of degree $\leqslant n-1$ and h_2 is a multiple of p. This decomposition is unique. Denote by F the linear span of $e_0, e_1,\dots,e_{n-1}$, and by G the n-dimensional space $\operatorname{Ker}\tilde{p}(S)$. It follows that F and G are dual to each other and that $G^{\perp}$ coincides with the subspace of all multiples of p.

If $\alpha_1,\dots,\alpha_n$ are all different, the space G is generated by the evaluation functionals $e(\alpha_1),\dots,e(\alpha_n)$ and the decomposition $h=h_1+h_2$ may be characterized by the interpolation property of h_1: h_1 is the (only) polynomial of degree $\leqslant n-1$ which assumes the same value as h at the points $\alpha_1, \ldots, \alpha_n$. Then h_2 is divisible by p, since it assumes the value zero at each point α_i. In this case the duality of F and G is easily established, since

$$\det(e_r, e(\alpha_s)) = \det \alpha_s^r$$

is the Vandermonde determinant of the roots $\alpha_1, \ldots, \alpha_n$.

If the polynomial p has multiple roots, then a basis for G may be constructed as follows. Given a root α of multiplicity m, we take the (obviously linearly independent) functions

$$e(\alpha), e(\alpha)^2, \ldots, e(\alpha)^m;$$

the union of these sets if α ranges over all distinct roots of p forms a basis for G. The corresponding characterization of h_1 is then that h_1 is the (only) polynomial of degree $\leqslant n-1$ with the following property: for each root α of multiplicity m the values of the derivatives of h and h_1 at α coincide, i.e.,

$$h_1^{(k)}(\alpha) = h^{(k)}(\alpha) \qquad \text{for} \quad k = 0, 1, \ldots, m-1.$$

Denote by $g = (g_0, \ldots, g_{n-1})$ the basis of G for which $G(e, g) = 1$, and by Π the projection onto F for which $\operatorname{Ker} \Pi = G^{\perp}$, so that

$$\Pi f = \sum (f, g_j) e_j$$

for every $f \in H^2$.

If A is an operator for which $p(A) = 0$, we have

$$f(A) = \sum_{j=0}^{n-1} (f, g_j) A^j,$$

since $f - \Pi f$ is a multiple of p.

Let us clear up the connection of the functions g_j with the matrix T^{∞}. The basic property of T^{∞} is the possibility of expressing all nonnegative powers of an n-dimensional operator in terms of the first $n-1$ powers. Thus

$$A^r = \sum_{k=0}^{n-1} t_{rk} A^k$$

for every operator A for which $p(A)=0$. In particular

$$\alpha^r = \sum_{k=0}^{n-1} t_{rk}\alpha^k$$

for every α in the spectrum of A. Suppose for a moment that all the roots of p are distinct. The above equation says then that $z^r - \sum_{k=0}^{n-1} t_{rk} z^k$ is divisible by p, so that $\Pi e_r = \sum_{k=0}^{n-1} t_{rk} e_k$ for every $r = 0,1,\ldots$. The coefficients of g_k may thus be obtained as follows:

$$(g_k, e_r) = (e_r, g_k)^* = (\Pi e_r, g_k)^* = t_{rk}^*,$$

so that $g_k = \tilde{c}_k$ if c_k stands for the kth column of T^∞.

We have thus proved that $\Pi = \Pi(F, G^\perp)$ may be expressed as

$$\Pi x = \sum_{j=0}^{n-1} (x, \tilde{c}_j) e_j.$$

To avoid continuity arguments, let us show how this relation may be proved even in the general case.

Let $P: H^2 \to C^n$ be the operator which assigns to each sequence $(x_0, x_1, \ldots) \in H^2$ the n-vector $Px = (x_0, x_1, \ldots, x_{n-1})^T$. The adjoint of P is thus the injection V of C^n into H^2,

$$V\big((x_0, \ldots, x_{n-1})^T\big) = (x_0, x_1, \ldots, x_{n-1}, 0, 0, \ldots).$$

Since the columns $c_0, \ldots, c_{n-1}$ of $T^\infty(p)$ are linearly independent, they generate $\operatorname{Ker} p(S)$. The composition $T^\infty(p)P$,

$$T^\infty(p)Px = \sum (x, e_j) c_j,$$

obviously coincides with $\Pi(\operatorname{Ker} p(S), F^\perp)$, so that its adjoint is

$$VT^\infty(p)^* = \Pi\big(F, \operatorname{Ker} p(S)^\perp\big),$$

$$VT^\infty(p)^* x = \sum (x, c_j) e_j.$$

Now $\operatorname{Ker} p(S)^\perp$ is nothing more than the set of all multiples of $\tilde{p}$. In order to obtain the projection onto F along the multiples of p it suffices to replace the elements of $T^{\infty}(p)$ by their conjugates, so that

$$\Pi = \Pi\left(F, \operatorname{Ker}\tilde{p}(S)^{\perp}\right) = VT^{\infty}(p)^{T},$$

$$\Pi x = \sum (x, \tilde{c}_j) e_j.$$

In this manner the system g_j, characterized by the postulates that $g_0, \ldots, g_{n-1}$ be orthogonal to $e_0, \ldots, e_{n-1}$ and $g_j \in G = \operatorname{Ker}\tilde{p}(S)$, coincides with $\tilde{c}$:

$$g_j = \tilde{c}_j.$$

The rest of this section will be devoted to the investigation of a formula important for the study of the matrix T^{∞}. This formula (see formula below) was established first in [4], where it was used to obtain an explicit expression for the coefficients t_{rk}.

Let us define a function $F(t, z)$ for $|t| < 1$ and $|z| < 1$ as the value at z of the function $\Pi e(t^*)$. We have thus

$$\Pi e(t^*) = \Pi \sum_{r=0}^{\infty} t^r e_r = \sum_{r=0}^{\infty} t^r \sum_{k=0}^{n-1} t_{rk} e_k$$
$$= \sum_{k=0}^{n-1} \left(\sum_{r=0}^{\infty} t_{rk} t^r \right) e_k = \sum_{k=0}^{n-1} f_k(t) e_k,$$

where $f_k = \tilde{g}_k$. Hence

$$F(t, z) = \sum_{k=0}^{n-1} f_k(t) z^k.$$

In a somewhat less precise form this may be written as

$$\Pi \frac{1}{1 - tz} = F(t, z) = \sum_{k=0}^{n-1} f_k(t) z^k$$

In [4] the following explicit expression for F was obtained:

$$F(t, z) = \frac{1}{1 - tz}\left(1 - t^n \frac{p(z)}{q(t)}\right),$$

where q stands for the polynomial p_1, so that $q(t) = t^n p(1/t)$.

In what follows we present a series of propositions, some of which are of independent interest; they provide another proof of the formula above as well as some other interesting consequences. Some of the already known results appear thus in a different light, which provides more insight into the matter.

PROPOSITION 2.1. *Let p be a polynomial of degree n, and let q be the polynomial $q(t) = t^n p(1/t)$. Then*

$$S^n q(S^*) = p(S).$$

Proof. Since $SS^* = I$, we have

$$S^n q(S^*) = S^n \sum_{j=0}^{n} a_{n-j} S^{*j} = \sum_{j=0}^{n} a_{n-j} S^{n-j} = p(S). \qquad \blacksquare$$

An immediate consequence of this is the following corollary.

PROPOSITION 2.2. *Suppose the sequence $x = x_0, x_1, \dots$ satisfies the recurrence relation*

$$\sum_{j=0}^{n} a_j x_{r+j} = 0 \qquad \text{for all} \quad r = 0, 1, \dots .$$

Then the product $q(t)\Sigma_{s=0}^{\infty} x_s t^s$ is a polynomial of degree $\leqslant n-1$.

Proof. The element x satisfies $p(S)x = 0$. Since the product $q(t)\Sigma x_s t^s$ is nothing more than $q(S^*)x$, it is annihilated by S^n and is, accordingly, a polynomial of degree $\leqslant n-1$. A less sophisticated computational proof goes as follows. The coefficient of t^k in the product $q(t)x(t)$ is $\Sigma a_j x_s$ for all pairs j, s such that $n - j + s = k$, $0 \leqslant j \leqslant n$, and $s \geqslant 0$. The second constraint means that $k - n + j$ should be nonnegative and is superfluous if $k \geqslant n$. In this case all indices $0 \leqslant j \leqslant n$ are admissible and the sum is zero. $\blacksquare$

The function $F(t, z) = \Sigma_{k=0}^{n-1} f_k(t) z^k$ is a polynomial of degree $\leqslant n-1$ in z. The functions f_k belong to $\operatorname{Ker} p(S)$, since $f_k = \tilde{g}_k$ and $g_k \in \operatorname{Ker} \tilde{p}(S)$. It follows from Proposition 2.2 that $q(t)F(t, z)$ will be a polynomial in t of degree not exceeding $n-1$. Hence

$$q(t)F(t, z) = \sum_{j=0}^{n-1} w_j(t) z^j$$

for some polynomials w_j of degree $\leqslant n-1$. At the same time it follows from the definition of F that

$$\frac{1}{1-tz} - F(t,z)$$

is a multiple of $p(z)$, so that

$$\begin{aligned} q(t) - (1-tz)\sum w_j z^j &= q(t) - (1-tz)q(t)F(t,z) \\ &= (1-tz)q(t)\left(\frac{1}{1-tz} - F(t,z)\right) \end{aligned}$$

is a multiple of $p(z)$. We can thus write

$$q(t) - (1-tz)\sum w_j z^j = p(z)m(t,z).$$

Write m in the form

$$m(t,z) = m_0(t) + m_1(t)z + m_2(t)z^2 + \cdots .$$

As a function of z the left-hand side of the above identity is a polynomial of degree $\leqslant n$; it follows that $m_1 = m_2 = \cdots = 0$, whence

$$q(t) - (1-tz)\sum w_j z^j = p(z)a(t)$$

for some polynomial $a(t)$ of degree $\leqslant n$. Now it is easy to see that there exists exactly one polynomial $a(t)$ for which

$$q(t) - a(t)p(z)$$

is divisible by $1-tz$. This is the polynomial $a(t) = t^n$. Indeed,

$$\begin{aligned} q(t) - t^n p(z) &= \sum_{r=0}^{n} (a_r t^{n-r} - a_r z^r t^n) \\ &= \sum_{r=1}^{n} a_r t^{n-r}(1 - z^r t^r), \end{aligned}$$

so that $q(t)-t^n p(z)$ is divisible by $1-tz$. On the other hand, if $a(t)$ is a polynomial for which $q(t)-p(z)a(t)$ is divisible by $1-tz$, the expression

$$t^n p\left(\frac{1}{t}\right)-a(t)p(z)$$

is a multiple of $1-tz$ and hence becomes zero if we set $z=1/t$. It follows that

$$[t^n-a(t)]p\left(\frac{1}{t}\right)=0$$

and $a(t)=t^n$.

Summing up, we have

$$(1-tz)\sum w_j z^j=q(t)-t^n p(z),$$

whence $(1-tz)q(t)F(t,z)=q(t)-t^n p(z)$.

A somewhat more explicit form of the polynomial $\Sigma w_j(t)z^j$ may be obtained as follows. Denote by r the polynomial for which

$$p(y)-p(z)=(y-z)r(y,z),$$

so that r is of degree $\leqslant n-1$ both in y and z. Hence

$$\begin{aligned} q(t)-t^n p(z) &= t^n\left[p\left(\frac{1}{t}\right)-p(z)\right] \\ &= t^n\left(\frac{1}{t}-z\right)r\left(\frac{1}{t},z\right)=(1-tz)t^{n-1}r\left(\frac{1}{t},z\right), \end{aligned}$$

so that

$$\sum w_j(t)z^j=t^{n-1}r\left(\frac{1}{t},z\right).$$

3. THE LYAPUNOV EQUATION

Let us sketch briefly how the construction of orthogonal systems described at the end of Section 1 may be used with advantage to prove explicit formulae for the solution of Lyapunov equations.

First of all, it follows from the second property of $T^{\infty}(p)$ that postmultiplication by $C(p)$ amounts to the same as shifting $T^{\infty}(p)$ one row up. In other words, interpreting $T^{\infty}(p)$ as a mapping of C^n into H^2, the following relation holds:

$$T^{\infty}(p)C(p) = ST^{\infty}(p).$$

Given two polynomials a, b of degree n, we have

$$\begin{aligned} C(a)^*T^{\infty}(a)^*T^{\infty}(b)C(b) &= T^{\infty}(a)^*S^*ST^{\infty}(b) \\ &= T^{\infty}(a)^*(I - E_0)T^{\infty}(b) = T^{\infty}(a)^*T^{\infty}(b) - E, \end{aligned}$$

where we have denoted by E_0 the projection of H^2 onto the constant functions and by E the matrix

$$E = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

This fact may be used to give an explicit expression for the solution of the Lyapunov equation

$$X - C(a)^*XC(b) = E.$$

Indeed, the identity above shows that $X = T^{\infty}(a)^*T^{\infty}(b)$ is the solution, and it follows from the first property of the infinite companion matrix that X is nothing more than the Gram matrix of the bases

$$c(b) = (c_0(b),\dots,c_{n-1}(b)) \quad \text{and} \quad c(a) = (c_0(a),\dots,c_{n-1}(a)).$$

Although explicit expressions for the functions c_j are known, much neater formulae for the solution of the Lyapunov equation may be obtained using Lemma 1.5. Set

$$F = \operatorname{Ker} a(S), \qquad G = \operatorname{Ker} b(S),$$

$$f_i^{\perp} = \Pi(G, F^{\perp})e_i, \qquad g_i^{\perp} = \Pi(F, G^{\perp})e_i.$$

Then

$$G(c(b),c(a))=G(f^{\perp},g^{\perp})^{-1}$$

and

$$G(f^{\perp},g^{\perp})_{ik}=(f_k^{\perp},g_i^{\perp})=(\Pi(G,F^{\perp})e_k,\Pi(F,G^{\perp})e_i)$$
$$=(\Pi(G,F^{\perp})e_k,e_i).$$

Now we shall use Lemma 1.6 to show that

$$\Pi(G,F^{\perp})=I-b_0(S)^{*-1}a(S)^*b(S)a_0(S)^{-1}.$$

Indeed, we have for $A=a(S)$ and $B=b(S)$

$$\Pi(G,F^{\perp})=I-\Pi(F^{\perp},G)$$
$$=I-A^*YB=I-a(S)^*[b(S)a(S)^*]^{-1}b(S),$$

and it suffices to observe that

$$b(S)a(S)^*=a_0(S)b_0(S)^*$$

(see the last formula on p. 370 in [6]). We have thus

$$G(f^{\perp},g^{\perp})=(\Pi(G,F^{\perp})e_k,e_i)$$
$$=([I-b_0(S)^{*-1}a(S)^*b(S)a_0(S)^{-1}]e_k,e_i)$$
$$=((I-b_0(S_n)^{*-1}a(S_n)^*b(S_n)a_0(S_n)^{-1})_{ik}),$$

so that

$$X=G(f^{\perp},g^{\perp})^{-1}=[I-b_0(S_n)^{*-1}a(S_n)^*b(S_n)a_0(S_n)^{-1}]^{-1}.$$

This formula was obtained first and the proof subsequently greatly simplified [8] by N. J. Young.

4. THE INTERPOLATION FORMULA

As an application of the preceding considerations we intend to derive the classical interpolation formula for functions $m \in H^2$. An explicit expression for Πm may be obtained using the following identity:

$$F(y, z) = \frac{1}{1 - yz}\left(1 - \frac{p(z)}{p(1/y)}\right).$$

For convenience let us recall that $\Pi = \Pi(F, \operatorname{Ker} \tilde{p}(S)^{\perp})$ and that $F(y, z) = (\Pi e(y^*), e(z))$. In particular, if $|y| = 1$ then

$$F(y^*, z) = \frac{1}{1 - y^*z}\left(1 - \frac{p(z)}{p(y)}\right);$$

using this, we obtain, for $y = e^{it}$,

$$\begin{aligned}
\Pi m &= \sum (m, g_k) z^k \\
&= \sum \frac{1}{2\pi} \int m(y)[g_k(y)]^* z^k \, dt \\
&= \frac{1}{2\pi} \int m(y) \sum f_k(y^*) z^k \, dt \\
&= \frac{1}{2\pi} \int m(y) F(y^*, z) \, dt \\
&= \frac{1}{2\pi i} \int m(y) F(y^*, z) \frac{dy}{y} \\
&= \frac{1}{2\pi i} \int m(y) \frac{1}{1 - y^*z}\left(1 - \frac{p(z)}{p(y)}\right) \frac{dy}{y} \\
&= \frac{1}{2\pi i} \int \frac{m(y)}{y - z}\left(1 - \frac{p(z)}{p(y)}\right) dy.
\end{aligned}$$

REFERENCES

1 Z. Dostál, Polynomials of the eigenvalues and powers of matrices, *Comment. Math. Univ. Carolin.* 19:459–468 (1978).
2 V. Pták, Norms and the spectral radius of matrices, *Czechoslovak Math. J.* 87:555–557 (1962).
3 V. Pták, Rayon spectral, norme des itérés d'un opérateur et exposant critique, *C. R. Acad. Sci. Paris* 265:257–259 (1967).
4 V. Pták, Spectral radius, norms of iterates and the critical exponent, *Linear Algebra Appl.* 1:245–260 (1968).
5 V. Pták, An infinite companion matrix, *Comment. Math. Univ. Carolin.* 19:447–458 (1978).
6 V. Pták and N. J. Young, Functions of operators and the spectral radius, *Linear Algebra Appl.* 29:357–392 (1980).
7 B. Sz.-Nagy, Sur la norme des fonctions de certains opérateurs, *Acta Math. Acad. Sci. Hungar.* 20:331–334 (1969).
8 N. J. Young, The solution of Lyapunov matrix equations by a geometrical method, *Math. Proc. Royal Soc. Edinburgh* 86A:347–354 (1980).

Received 27 April 1981; revised 12 May 1982