W. G. Tuel, Jr.

Computer Algorithm for Spectral Factorization of Rational Matrices

Abstract: An algorithm is derived for the numerical spectral factorization of matrices arising in optimal filter design. The method uses a bilinear transformation to convert the factorization problem into a stable nonlinear difference equation of the Riccati type. The computer solution of several examples is presented to illustrate the technique.

I. Introduction

In the determination of optimal filters for stationary random processes, the filter transfer function is given as the solution of a Wiener-Hopf integral equation.[1–3] It has long been recognized that Wiener's method of spectral factorization is the most elegant means of solution of this equation. Recently, several authors have considered the solution of multivariable filtering problems by this technique.[4–11,14,15] The chief computational stumbling block is the spectral factorization of a rational matrix.

Youla[5] suggested a method based on his constructive proof that the desired factorization does exist. One transforms the given matrix into its Smith canonical (diagonal) form by multiplying by elementary matrices, factors the diagonal polynomial elements, and then uses an ingenious but intricate sequence of matrix manipulations to obtain the final result. Davis[8] simplified the problem somewhat by showing that the matrix to be factored could be reduced in steps to a constant matrix by appropriately chosen transformations. The product of these transformations is the desired factor. As in Youla's algorithm, however, one must factor a scalar polynomial and perform several matrix transformations.

Recently, Riddle and Anderson[11] have suggested that the difficult step is factoring the scalar polynomials and have pointed out that, since the roots themselves are not needed, methods for obtaining the factors directly can be employed. They describe a technique in which one obtains the coefficients of the desired factor from a set of nonlinear algebraic equations, the iterative solution of which is said to be convergent.

However, although many techniques exist, all are similar in that the matrix to be factored must be manipulated in such a way that scalar polynomials may be extracted and factored. The decision problems inherent in such manipulations (e.g., determination of linear dependence) make such algorithms difficult to program for a digital computer.

This paper describes a straightforward algorithm for spectral factorization of matrices which avoids the difficulties described above. The matrix to be factored is converted directly into a nonlinear difference equation of the Riccati type, which converges to a stationary solution. The desired factor is obtained easily from the solution. In the scalar case, the algorithm closely resembles the one proposed by Riddle and Anderson.

- *Problem statement and formal solution*

The optimal (minimum mean square error) filter transfer function $\phi(t)$ is the solution of the matrix Wiener-Hopf integral equation

$$\int_0^\infty \Gamma(t-\tau)\Phi(\tau)d\tau = \psi(t)\,, \qquad t > 0\,, \tag{1.1}$$

where $\Gamma(t)$ is an $r \times r$ covariance matrix, $\psi(t)$ a known column vector, and $\phi(t)$ the desired solution. It will be assumed throughout that the spectral density $G(s)$,

$$G(s) = \int_{-\infty}^\infty \Gamma(t)e^{-st}dt\,, \tag{1.2}$$

is an $r \times r$ rational matrix with the following properties:

A: 1) $G(s)$ is real, i.e., $\bar{G}(s) = G(\bar{s})$;
2) $G(s)$ is para-Hermitian, i.e., $G^T(-s) = G(s)$;
3) $G(s)$ is of normal rank r (almost everywhere);
4) $G(j\omega)$ is positive semidefinite for every finite ω.

Then there exists an $r \times r$ matrix $H(s)$ such that:

B: 1) $G(s) = H^T(-s)H(s)$; (1.3)

2) $H(s)$ is real, i.e., $H(s) = \bar{H}(s)$;

3) $H(s)$ and $H^{-1}(s)$ are both rational and analytic in the right-half plane, i.e., all their poles and zeroes lie in the left-hand plane.

Once $H(s)$ has been found, the solution transfer function is

$$\Phi(s) = \int_0^\infty \phi(t)e^{-st}dt$$
$$= H^{-1}(s)\{[H^T(-s)]^{-1}\Psi(s)\}_+ \quad (1.4)$$

where $\Psi(s) = \int_0^\infty \psi(t)e^{-st}dt$

and $\{\ \ \}_+$ denotes that part of the partial fraction expansion of $[H^T(-s)]^{-1}\Psi(s)$ due to the poles of $\Psi(s)$.

$H^{-1}(s)$ may be evaluated by (for example) Gauss reduction, while the $\{\ \ \}_+$ term may be computed in the following way. Let

$$\Psi(s) = \frac{\Sigma(s)}{\sigma(s)} \text{ and } [H^T(-s)]^{-1} = \frac{\Theta(-s)}{\theta(-s)}, \quad \text{where}$$

$\Sigma(s)$, $\Theta(s)$ are matrix polynomials and $\sigma(s)$, $\theta(s)$ are scalar polynomials. Then

$$\{[H^T(-s)]^{-1}\Psi(s)\}_+ = \left\{\frac{\Theta(-s)\Sigma(s)}{\theta(-s)\sigma(s)}\right\}_+$$
$$= \left\{\frac{\Delta^-(s)}{\theta(-s)} + \frac{\Delta^+(s)}{\sigma(s)}\right\}_+ = \frac{\Delta^+(s)}{\sigma(s)}$$

where $\Delta^+(s)\theta(-s) + \Delta^-(s)\sigma(s) = \Theta(-s)\Sigma(s)$. (1.5)

Since $\Delta^{\pm}(s)$, $\Sigma(s)$ are polynomial vectors, each element of $\Delta^+(s)$ can be obtained by solving a (relatively small) set of simultaneous linear equations for the coefficients of the polynomials. Note also that it is never necessary to compute the roots of a polynomial to obtain the optimum transfer function.

- *Spectral factorization*

The principal problem, then, is the factorization expressed by (1.3). It is to this problem that the rest of the paper is devoted. The rational matrix $G(s)$ may always be written

$$G(s) = \frac{\tilde{G}(s)}{g(s)}, \quad (1.6)$$

where $\tilde{G}(s)$, $g(s)$ are respectively matrix and scalar polynomials. If both are factored into products of polynomials, i.e.,

$$\tilde{G}(s) = \tilde{H}^T(-s)\tilde{H}(s) \quad (1.7)$$
$$\text{and} \quad g(s) = h(-s)h(s),$$

where $\tilde{H}(s)$, $h(s)$ are polynomials with their roots in the left-hand plane, then

$$H(s) = \tilde{H}(s)/h(s) \quad (1.8)$$

is the desired factor of $G(s)$. Thus there is no loss of generality in discussing the spectral factorization of polynomial matrices, since any rational matrix may be factored by applying the algorithm separately to numerator and denominator.

II. Factorization of polynomial matrices

$$\text{Let } G(s) = \left\{\sum_{k=0}^{n} g_{ijk}s^k\right\} \quad (2.1)$$

be an $r \times r$ polynomial matrix with properties (**A1–A4**) of Section I. The problem is to find an $r \times r$ polynomial matrix

$$H(s) = \left\{\sum_{k=0}^{m} h_{ijk}s^k\right\} \quad (m = n/2) \quad (2.2)$$

with properties (**B1–B3**).

Since a digital computer is to be used for the calculations, it is convenient to map the continuous variable s onto the discrete variable z by the transformation pair

$$z = \frac{1+s}{1-s}, \quad (2.3a)$$

$$s = \frac{z-1}{z+1}, \quad (2.3b)$$

solve the resulting discrete factorization problem, then map the solution back into the continuous plane. The algorithm for discrete factorization, of course, may be used directly if the original problem is filtering of discrete-time processes.[9]

The details of the mapping (2.3) are discussed below. The discrete factorization algorithm is derived in the next section. Let m_i denote the highest power of s^2 appearing in the i^{th} diagonal element of $G(s)$, i.e.,

$$G_{ii}(s) = \sum_{k=0}^{2m_i} g_{iik}s^k, \qquad i = 1, 2, \cdots, r. \quad (2.4)$$

Then, by reason of the positive semidefiniteness of $G(s)$, the highest power of s that may occur in the i, j^{th} off-diagonal element is $m_i + m_j$, that is,

$$G_{ij}(s) = \sum_{k=0}^{m_i+m_j} g_{ijk}s^k. \quad (2.5)$$

Substituting (2.3b) into (2.5) and dividing by z^{m_j}, one gets

$$G_{ij}\left(\frac{z-1}{z+1}\right) = \frac{1}{(1+z)^{m_i}}$$
$$\times\left\{\sum_{k=0}^{m_i+m_j} g_{ijk}(z-1)^k(z+1)^{m_i+m_j-k}(z)^{-m_j}\right\}$$
$$\times\frac{1}{(1+z^{-1})^{m_j}}$$
$$= \frac{1}{(1+z)^{m_i}}\left\{\sum_{k=-m_i}^{m_j} c_{ijk}z^{-k}\right\}\frac{1}{(1+z^{-1})^{m_j}}. \quad (2.6)$$

Thus, the complete matrix may be written:

$$G\left(\frac{z-1}{z+1}\right) = \begin{bmatrix} (1+z)^{-m_1} & & \\ & \ddots & \\ & & (1+z)^{-m_r} \end{bmatrix}$$
$$\times\, C(z) \begin{bmatrix} (1+z^{-1})^{-m_1} & & \\ & \ddots & \\ & & (1+z^{-1})^{-m_r} \end{bmatrix} \quad (2.7)$$

where the i, j^{th} element of $C(z)$

$$C_{ij}(z) = \sum_{k=-m_i}^{m_j} c_{ijk} z^{-k} = C_{ji}(z^{-1}) \quad (2.8)$$

since $G(s)$ is para-Hermitian. It can be verified that $C(z)$ satisfies the conditions:

C: 1) $C(z)$ is real;
2) $C(z)$ is para-Hermitian, i.e., $C(z) = C^T(1/z)$;
3) $C(z)$ is of normal rank r (almost everywhere);
4) $C(e^{i\Theta})$ is positive semidefinite for all Θ, $0 \leq \Theta < 2\pi$;
5) The matrix $\{c_{ij0}\}$ is positive definite, since

$$\{c_{ij0}\} = \frac{1}{2\pi}\int_0^{2\pi} C(e^{i\Theta})\,d\Theta\,.$$

Thus* $C(z)$ has a factorization

$$C(z) = D^T\left(\frac{1}{z}\right)D(z) \quad (2.9)$$

where

$$D_{ij}(z) = \sum_{k=0}^{m_j} d_{ijk} z^{-k} \quad (2.10)$$

is of normal rank r and has all its roots inside or on the unit circle.

Under the inverse transformation (2.3a) the i, j^{th} element of the product

$$D(z) \begin{bmatrix} (1+z^{-1})^{-m_1} & & \\ & \ddots & \\ & & (1+z^{-1})^{-m_r} \end{bmatrix}$$

becomes

$$H_{ij}(s) = \frac{1}{2^{m_j}} \sum_{k=0}^{m_j} d_{ijk}(1-s)^k(1+s)^{m_j-k}\,. \quad (2.11)$$

Thus, from (2.7), (2.9), and (2.11),

$$G(s) = H^T(-s)H(s) \quad (2.12)$$

as desired.

*(One can establish this by adapting Youla's proof[8] for the continuous problem.)

III. Factorization of discrete para-Hermitian matrices

The aim of this section is to reduce the factorization of the para-Hermitian polynomial matrix $C(z)$ to the solution of a set of nonlinear algebraic equations. An iterative method for the solution of these equations is proposed. In Section IV it will be shown that this iteration does, in fact, converge to the desired solution.

The following quantities are used throughout this section and are collected here for ease of reference:

Define:

$$M = \sum_{i=1}^{r} m_i \quad (3.1)$$

$$\mathbf{c}_{ij}^T \equiv [c_{ijm_j} c_{ijm_j-1} \cdots c_{ij1}]\,, \quad (3.2)$$

$$\mathbf{p}_i^T(z) \equiv [z^{-m_i} z^{-(m_i-1)} \cdots z^{-1}]\quad i = 1, 2, \cdots r\,, \quad (3.3)$$

$$P^T(z) \equiv \begin{bmatrix} \mathbf{p}_1(z) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{p}_2(z) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{p}_r(z) \end{bmatrix} \quad (3.4)$$

of dimension $M \times r$,

$$Q_{12} \equiv \begin{bmatrix} \mathbf{c}_{11} & \mathbf{c}_{21} & \cdots & \mathbf{c}_{r1} \\ \mathbf{c}_{12} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathbf{c}_{1r} & \cdots & & \mathbf{c}_{rr} \end{bmatrix} \quad (3.5)$$

of dimension $M \times r$,

$$Q_{22} \equiv \begin{bmatrix} c_{110} & \cdots & c_{1r0} \\ c_{120} & \ddots & \vdots \\ \vdots & & \vdots \\ c_{1r0} & \cdots & c_{rr0} \end{bmatrix} \quad (3.6)$$

of dimension $r \times r$,

$$Z_i \equiv \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & & 1 & \ddots & \vdots \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix} \quad (3.7)$$

of dimension $m_i \times m_i$,

$$\mathbf{e}_i \equiv \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (3.8)$$

of dimension $m_i \times 1$,

$$A_{11} \equiv \begin{bmatrix} Z_1 & 0 & \cdots & 0 \\ 0 & Z_2 & \ddots & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & Z_r \end{bmatrix} \quad (3.9)$$

of dimension $M \times M$,

$$A_{12} \equiv \begin{bmatrix} e_1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & e_2 & \vdots & \cdot & \cdot & \vdots \\ \vdots & & & \cdot & & \\ \vdots & & & & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & e_r \end{bmatrix} \tag{3.10}$$

of dimension $M \times r$.
It is easily verified that

$$C(z) = Q_{22} + Q_{12}^T P(z) + P^T\left(\frac{1}{z}\right)Q_{12} . \tag{3.11}$$

Theorem:

Let $C(z) = Q_{22} + Q_{12}^T P(z) + P^T(1/z)Q_{12}$ (3.11)

have a factorization

$$C(z) = D^T(1/z)D(z) , \tag{2.9}$$

where $D(z)$ is a polynomial in z^{-1} with its roots inside or on the unit circle. Then $D(z)$ is expressible as

a) $D(z) = W_2^T + W_1^T P(z)$ (3.12)

where W_2 is an $r \times r$ positive definite triangular matrix and W_1 is an $M \times r$ matrix.

b) W_1, W_2 are obtained from the equations

$$W_2 W_2^T = X_{22} \tag{3.13}$$

$$W_1 W_2^T = X_{12} , \tag{3.14}$$

where X_{12}, X_{22}, X_{11}, Y satisfy

$$X_{12} = Q_{12} + A_{11}^T Y A_{12} \tag{3.15}$$

$$X_{22} = Q_{22} + A_{12}^T Y A_{12} \tag{3.16}$$

$$Y = A_{11}^T Y A_{11} - X_{12} X_{22}^{-1} X_{12}^T . \tag{3.17}$$

Comment: Equations (3.15)–(3.17) are the crucial set. Once X_{12}, X_{22} have been obtained it is relatively easy to compute W_1, W_2, especially since W_2 is triangular.

Proof:

a) It is easily verified that any polynomial matrix factor in z^{-1} of normal rank r may be represented in the form

$$U(z) = U_2^T + U_1^T P(z) , \tag{3.18}$$

where U_2^T is nonsingular but not necessarily positive definite or triangular. Let W_2 be the positive definite triangular square root of $U_2U_2^T$, i.e. $W_2W_2^T = U_2U_2^T$. Then $U_2 = W_2(W_2^{-1}U_2)$, so that the polynomial $U(z)$ may be written

$$\begin{aligned} U(z) &= [W_2^{-1}U_2]^T\{W_2^T + W_2^T(U_2^T)^{-1}U_1^T P(z)\} \\ &= [W_2^{-1}U_2]^T\{W_2^T + W_1^T P(z)\} \\ &= [W_2^{-1}U_2]^T D(z) . \end{aligned}$$

Now two factorizations are unique only to within an arbitrary orthogonal matrix, i.e., any two factors $U_1(z)$, $U_2(z)$ are related by $U_1(z) = VU_2(z)$, where V is an orthogonal matrix. Since $[W_2^{-1}U_2]$ is orthogonal, $D(z)$ is also a factorization of $C(z)$.

b) Equation (2.9) may be expressed in terms of W_1 and W_2 as

$$\begin{aligned} C(z) = {} & W_2 W_2^T + P^T(1/z) W_1 W_2^T + W_2 W_1^T P(z) \\ & + P^T(1/z) W_1 W_1^T P(z) . \end{aligned} \tag{3.19}$$

Define

$$X_{22} = W_2 W_2^T . \tag{3.13}$$

$$X_{12} = W_1 W_2^T . \tag{3.14}$$

Since $W_1W_1^T = (W_1W_2^T)(W_2W_2^T)^{-1}(W_2W_1^T)$, Eq. (3.19) may be written

$$\begin{aligned} C(z) = {} & X_{22} + P^T(1/z) X_{12} + X_{12}^T P(z) \\ & + P^T(1/z) X_{12} X_{22}^{-1} X_{12}^T P(z) . \end{aligned}$$

Let Y satisfy (3.17) and substitute into the last term. One gets

$$\begin{aligned} C(z) = {} & X_{22} + P^T(1/z) X_{12} \\ & + X_{12}^T P(z) - P^T(1/z) Y P(z) \\ & + P^T(1/z) A_{11} Y A_{11} P(z) . \end{aligned} \tag{3.20}$$

Now it is easily shown that

$$P(z) = z^{-1}[A_{11}P(z) + A_{12}] . \tag{3.21}$$

Using this property, the term

$$\begin{aligned} P^T(1/z) Y P(z) = {} & P^T(1/z) A_{11}^T Y A_{11} P(z) \\ & + P^T(1/z) A_{11}^T Y A_{12} \\ & + A_{12}^T Y A_{11}^T P(z) \\ & + A_{12}^T Y A_{12} . \end{aligned} \tag{3.22}$$

Substituting (3.22) into (3.20) and collecting terms, one sees that

$$\begin{aligned} C(z) = {} & [X_{22} - A_{12}^T Y A_{12}] \\ & + P^T(1/z)[X_{12} - A_{11}^T Y A_{12}] \\ & + [X_{12} - A_{11} Y A_{12}]^T P(z) . \end{aligned} \tag{3.23}$$

Since (3.11) and (3.23) must be identical for all z,

$$X_{22} = Q_{22} + A_{12}^T Y A_{12} \tag{3.16}$$

$$X_{12} = Q_{12} + A_{11}^T Y A_{12} \tag{3.15}$$

This completes the set of equations required.

To solve (3.15)–(3.17), it is natural to investigate the iteration

$$\begin{aligned} & X_{12}(k+1) \\ & = Q_{12} + A_{11}^T Y(k) A_{12} \qquad X_{12}(0) = Q_{12} \end{aligned} \tag{3.24}$$

$$X_{22}(k+1) = Q_{22} + A_{12}^T Y(k) A_{12} \qquad X_{22}(0) = Q_{22} \tag{3.25}$$

$$X_{11}(k+1) = \quad + A_{11}^T Y(k) A_{11} \qquad X_{11}(0) = 0 \tag{3.26}$$

$$Y(k) = X_{11}(k) - X_{12}(k) X_{22}^{-1}(k) X_{12}^T . \tag{3.27}$$

This difference equation is the discrete matrix Riccati equation that would have been obtained had one solved the transformed discrete filtering problem as a Kalman filtering problem.[12] It is demonstrated in Section IV that the Riccati equation converges to the desired solution.

IV. Convergence of the algorithm

In the previous section the discrete matrix Riccati equation was derived as a purely algebraic algorithm for spectral factorization of discrete para-Hermitian matrices. In this section it will be shown (under the minor restriction that no roots lie exactly on the unit circle) that the algorithm converges; moreover, the convergence is monotonic, i.e., the difference between two successive matrix iterates is positive semidefinite.

It is advantageous to introduce some new notation at this point. Define the $(M + r) \times (M + r)$ matrices

$$A \equiv \begin{bmatrix} A_{11} & A_{12} \\ 0 & 0 \end{bmatrix} \tag{4.1}$$

$$X \equiv \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^T & X_{22} \end{bmatrix} \tag{4.2}$$

$$Q \equiv \begin{bmatrix} 0 & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix} \tag{4.3}$$

and the $r \times (M + r)$ matrices

$$B^T \equiv [0 \quad I] \tag{4.4}$$

$$K \equiv (B^T X B)^{-1} B^T X A \tag{4.5}$$

$$= (X_{22}^{-1} X_{12}^T,\ I) A . \tag{4.6}$$

Using these symbols Eqs. (3.24) − (3.27) may be compactly written

$$X(k+1) = Q + [A - BK(k)]^T X(k) [A - BK(k)] \tag{4.7}$$

$$X(0) = Q .$$

Let X^* denote the steady-state solution matrix of (4.7),

$$X^* = Q + (A - BK^*)^T X^* (A - BK^*) , \tag{4.8}$$

which may also be written

$$X^* = Q + A^T \begin{bmatrix} Y^* & 0 \\ 0 & 0 \end{bmatrix} A , \tag{4.9}$$

where

$$Y^* = - \sum_{j=0}^{m} (A_{11}^j)^T W_1 W_1^T (A_{11}^j) \tag{4.10}$$

from (3.13)–(3.17).

The chief result necessary to prove monotonic convergence is given in the following theorem paraphrased from one proven by Nishimura.[13]

Theorem (Nishimura):
Let $X_1(k)$, $X_2(k)$ be any two solutions of the difference equation

$$X(k+1) = Q + [A - BK(k)]^T X(k) [A - BK(k)] \tag{4.11}$$

$$K(k) = [B^T X(k) B]^{-1} B^T X(k) A . \tag{4.12}$$

Define $E(k) \triangleq X_1(k) - X_2(k)$ and assume $E(0)$ is positive semidefinite. Then

a) $E(k+1) = [A - BK_2(k)]^T \times \{E(k) - E(k) B [B^T X_1(k) B]^{-1} B^T E(k)\} \times [A - BK_2(k)]$

b) The expression $E(k) - E(k)B(B^T X_1(k) B)^{-1} B^T E(k)$ is positive semidefinite.

c) Hence, $E(k)$ is positive semidefinite for all k.

The following lemma is also necessary.

Lemma: The nonzero eigenvalues of the matrix $(A - BK^*)$ coincide with the nonzero roots of the polynomial matrix factor

$$D(z) = W_2^T + W_1^T P(z) \text{ in } (3.12) .$$

Proof: From (4.5), $K^* = [L^T, I]A$, (4.14)
where

$$L^T = (W_2^T)^{-1} (W_1^T) . \tag{4.15}$$

Using (4.14), (4.1) and (4.4), the matrix $A - BK^*$ becomes

$$A - BK^* = \begin{bmatrix} A_{11} & A_{12} \\ -L^T A_{11} & -L^T A_{12} \end{bmatrix} .$$

The eigenvalues are obtained from

$$\det [\lambda I - (A - BK^*)] = \det \begin{bmatrix} \lambda I - A_{11} & -A_{12} \\ L^T A_{11} & \lambda I + L^T A_{12} \end{bmatrix} = 0 ,$$

which may be manipulated into

$$\lambda^r \det [\lambda I - A_{11}] \det [(W_2^T)^{-1}] \times \det [W_2^T + W_1^T (I - A_{11} \lambda^{-1})^{-1} \lambda^{-1} A_{12}] = 0 . \tag{4.16}$$

Now the eigenvalues of A_{11} are all zero, and the expression

$$(I - A_{11}\lambda^{-1})^{-1}\lambda^{-1}A_{12} = P(\lambda)$$

arises from (3.21). Thus, the nonzero eigenvalues of $A - BK^*$ are obtained from the equation

$$\det\,[W_2^T + W_1^T P(\lambda)] = 0\,. \qquad \text{Q.E.D.}$$

Now, applying Nishimura's theorem, let $X_1(k)$ be the solution of Eq. (4.7), and $X_2(k) = X^*$, the steady-state solution (4.8). Then

$$E(0) = X_1(0) - X_2(0)$$
$$= Q - \left[Q + A^T\begin{pmatrix} Y^* & 0 \\ 0 & 0 \end{pmatrix} A\right]$$

is positive semidefinite from (4.10). Hence, $E(k)$ is positive semidefinite for all k. Since the bracket expression in (4.13) is positive semidefinite,

$$E(k+1) \leq (A - BK^*)^T E(k)(A - BK^*)\,. \qquad (4.17)$$

Thus

$$E(k) \leq [(A - BK^*)^k]^T E(0)[(A - BK^*)^k] \quad \text{for all } k\,. \qquad (4.18)$$

Now, under the slight restriction that the roots of $D(z)$ and thus also the eigenvalues of $(A - BK^*)$ are strictly within the unit circle, inequality (4.18) implies that $E(k)$ approaches 0 as k increases.

When one or more of the roots of $D(z)$ is on the unit circle, (4.18) gives no information. This circumstance occurs when the original polynomial matrix $G(s)$ is not strictly positive definite for all finite ω. Although convergence has not been proven for this case, in test cases the error decreases as $1/k$. In many instances the troublesome roots may be isolated (and convergence accelerated) by a preliminary factorization. See, e.g., sample problem 2).

To establish monotonicity, let

$$X_1(k) = X(k)$$
$$X_2(k) = X(k+1)\,.$$

Then

$$E(0) = X(0) - X(1)$$
$$= Q - \left[Q + A^T\begin{pmatrix} Y(0) & 0 \\ 0 & 0 \end{pmatrix} A\right]$$
$$= A^T\begin{bmatrix} Q_{12}Q_{22}^{-1}Q_{12}^T & 0 \\ 0 & 0 \end{bmatrix} A$$

is positive semidefinite. Thus $X(k) - X(k+1)$ is positive semidefinite by Nishimura's Theorem.

```
                    SPECTRAL FACTORIZATION PROGRAM

INPUT MATRIX G(S)

1 1   0.1250000E 03  -0.0000000E-38  -0.1000000E 01  -0.0000000E-38  -0.0000000E-38
1 2   0.1000000E 03  -0.0000000E-38  -0.0000000E-38  -0.0000000E-38  -0.0000000E-38
2 2   0.1250000E 03  -0.0000000E-38  -0.2600000E 02  -0.0000000E-38   0.1000000E 01

TRANSFORMED MATRIX C(Z)

1 1   0.2520000E 03   0.1240000E 03   0.0000000E-38
1 2   0.3000000E 03   0.3000000E 03   0.1000000E 03
2 1   0.3000000E 03   0.1000000E 03   0.0000000E-38
2 2   0.8080000E 03   0.4960000E 03   0.1000000E 03

ITERATION CONVERGED TO WITHIN 0.100E-04 IN    45 STEPS.  SOLUTION IS

1 1   0.1483389E 03   0.1240000E 03   0.0000000E-38
1 2   0.1182662E 03   0.1000000E 03   0.0000000E-38
2 1   0.1182662E 03   0.2162904E 03   0.1000000E 03
2 2   0.2912749E 03   0.3336053E 03   0.1000000E 03

THE MATRIX FACTOR D(Z) IS

1 1   0.1001595E 02   0.8326423E 01  -0.0000000E-38
1 2   0.0000000E-38   0.8070806E 01   0.5930245E 01
2 1   0.6929618E 01   0.5859337E 01   0.0000000E-38
2 2   0.1706678E 02   0.1954706E 02   0.5859337E 01

SOLUTION OF FACTORIZATION PROBLEM
G(S) = HT(-S)*H(S). H(S) IS

1 1   0.9171186E 01   0.8447636E 00   0.0000000E-38
1 2   0.3500263E 01  -0.2965122E 01  -0.5351402E 00
2 1   0.6394477E 01   0.5351404E 00   0.0000000E-38
2 2   0.1061829E 02   0.5603720E 01   0.8447634E 00

PRODUCT HT(-S)*H(S) IS

1 1   0.1250000E 03   0.0000000E-38  -0.1000001E 01  -0.0000000E-38   0.0000000E-38
1 2   0.1000000E 03   0.0000000E-38   0.1788139E-06  -0.1862645E-07   0.0000000E-38
2 2   0.1250000E 03   0.0000000E-38  -0.2600000E 02  -0.0000000E-38   0.1000000E 01
```

Figure 1 Computer output for sample problem 1.

V. Sample problems

To implement the factorization described in Sections II and III, a FORTRAN program package has been written. It is designed to factor polynomial matrices of dimension 6 × 6 or less, with order not exceeding 20 for each polynomial element. The total package requires approximately 10,000 words of storage on the IBM 7094. A complete source listing may be obtained by contacting the author.

The use of this program is illustrated in the following sample problems. The lines under the heading INPUT MATRIX $G(S)$ contain the polynomial coefficients of the upper right-half of the matrix to be factored, arranged in ascending powers of s. The next entries, headed TRANSFORMED MATRIX $C(Z)$, contain the coefficients of $C(z)$, beginning with z^0, and arranged in ascending powers of (z^{-1}). Coefficients of powers of z are redundant and not listed. The next line gives convergence information: whether the error criterion was met, and how many iterations were performed. Immediately following, solution information is written. Next, the discrete factor $D(z)$ is printed, labled THE MATRIX FACTOR $D(Z)$ IS. Again, the coefficients are arranged in ascending powers of (z^{-1}). The left-hand plane factor of $G(s)$ is written as SOLUTION OF FACTORIZATION PROBLEM . . . $H(S)$ IS. The coefficients are in ascending powers of s. Finally, the accuracy is checked by multiplying the factors together. The product is printed out for comparison with $G(s)$.

1) Consider the factorization of the polynomial matrix[8]

$$G(s) = \begin{bmatrix} 125 - s^2 & 100 \\ 100 & 125 - 26s^2 + s^4 \end{bmatrix}$$

The computer output is shown in Fig. 1.

The desired factor becomes

$$H(s) = \begin{bmatrix} 9.171186 & 3.500263 - 2.965122s \\ +.8447636s & -.5351402s^2 \\ & \\ 6.394477 & 10.61829 + 5.603720s \\ +.5351404s & +.8447634s^2 \end{bmatrix}.$$

2) Let[5]

$$G(s) = \begin{bmatrix} -s^2 & -s & 0 \\ s & 2 - s^2 & -s/2 \\ 0 & s/2 & -s^2 \end{bmatrix}.$$

This matrix is positive semidefinite only at the origin; hence, applied to this form the algorithm would converge very slowly. However, the roots at the origin can obviously be factored out by writing

$$G(s) = \begin{bmatrix} -s & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -s \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 - s^2 & -.5 \\ 0 & -.5 & 1 \end{bmatrix}$$

$$\times \begin{bmatrix} s & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & s \end{bmatrix}.$$

Then it is only necessary to factor the matrix

$$G(s) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 - s^2 & -.5 \\ 0 & -.5 & 1 \end{bmatrix}$$

The computer output is shown in Fig. 2. The factorization of $G(s)$ is then

$$H(s) = \begin{bmatrix} .8814124 & .4723475 - .4723475s & 0 \\ .4723475 & 1.235673 + .8814124s & 0 \\ 0 & -.5 & 1.0 \end{bmatrix}$$

$$\times \begin{bmatrix} s & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & s \end{bmatrix}$$

$$= \begin{bmatrix} .8814124s & .4723474 - .4723474s & 0 \\ .4723475s & 1.235673 + .8814124s & 0 \\ 0 & -.5 & s \end{bmatrix}.$$

```
                    SPECTRAL FACTORIZATION PROGRAM

INPUT MATRIX G(S)

1 1    0.1000000E 01   0.0000000E-38  -0.0000000E-38
1 2    0.1000000E 01   0.0000000E-38  -0.0000000E-38
1 3    0.0000000E-38   0.0000000E-38  -0.0000000E-38
2 2    0.2000000E 01  -0.0000000E-38  -0.1000000E 01
2 3   -0.5000000E 00  -0.0000000E-38  -0.0000000E-38
3 3    0.1000000E 01   0.0000000E-38  -0.0000000E-38

TRANSFORMED MATRIX C(Z)

1 1    0.1000000E 01   0.0000000E-38
1 2    0.1000000E 01   0.1000000E 01
1 3    0.0000000E-38   0.0000000E-38
2 1    0.1000000E 01   0.0000000E-38
2 2    0.6000000E 01   0.1000000E 01
2 3   -0.5000000E 00   0.0000000E-38
3 1    0.0000000E-38   0.0000000E-38
3 2   -0.5000000E 00  -0.5000000E 00
3 3    0.1000000E 01   0.0000000E-38

ITERATION CONVERGED TO WITHIN 0.100E-05 IN     4 STEPS.  SOLUTION IS

1 1    0.1000000E 01   0.0000000E-38
1 2    0.1000000E 01   0.0000000E-38
1 3    0.0000000E-38   0.0000000E-38
2 1    0.1000000E 01   0.1000000E 01
2 2    0.4732051E 01   0.1000000E 01
2 3   -0.5000000E 00  -0.5000000E 00
3 1    0.0000000E-38   0.0000000E-38
3 2   -0.5000000E 00   0.0000000E-38
3 3    0.1000000E 01   0.0000000E-38

THE MATRIX FACTOR D(Z) IS

1 1    0.8814124E 00  -0.0000000E-38
1 2    0.0000000E-38   0.9446950E 00
1 3    0.0000000E-38  -0.0000000E-38
2 1    0.4723475E 00   0.0000000E-38
2 2    0.2117085E 01   0.3542606E 00
2 3    0.0000000E-38   0.0000000E-38
3 1    0.0000000E-38   0.0000000E-38
3 2   -0.5000000E 00  -0.5000000E 00
3 3    0.1000000E 01   0.0000000E-38

SOLUTION OF FACTORIZATION PROBLEM
G(S) = HT(-S)*H(S). H(S) IS

1 1    0.8814124E 00   0.0000000E-38
1 2    0.4723475E 00  -0.4723475E 00
1 3    0.0000000E-38   0.0000000E-38
2 1    0.4723475E 00   0.0000000E-38
2 2    0.1235673E 01   0.8814124E 00
2 3    0.0000000E-38   0.0000000E-38
3 1    0.0000000E-38   0.0000000E-38
3 2   -0.5000000E 00   0.0000000E-38
3 3    0.1000000E 01   0.0000000E-38

PRODUCT HT(-S)*H(S) IS

1 1    0.1000000E 01  -0.0000000E-38  -0.0000000E-38
1 2    0.1000000E 01  -0.0000000E-38  -0.0000000E-38
1 3    0.0000000E-38  -0.0000000E-38  -0.0000000E-38
2 2    0.2000000E 01  -0.0000000E-38  -0.1000000E 01
2 3   -0.5000000E 00  -0.0000000E-38  -0.0000000E-38
3 3    0.1000000E 01  -0.0000000E-38  -0.0000000E-38
```

Figure 2 Computer output for sample problem 2.

Figure 3 Computer output for sample problem 3.

```
                    SPECTRAL FACTORIZATION PROGRAM

INPUT MATRIX G(S)

1 1    0.3000000E 01  -0.0000000E-38  -0.1000000E 01  -0.0000000E-38  -0.0000000E-38
      -0.0000000E-38  -0.0000000E-38  -0.0000000E-38  -0.0000000E-38
1 2   -0.0000000E-38  -0.0000000E-38  -0.0000000E-38  -0.0000000E-38   0.1000000E 01
      -0.0000000E-38  -0.0000000E-38  -0.0000000E-38  -0.0000000E-38
2 2    0.2000000E 01  -0.0000000E-38  -0.0000000E-38  -0.0000000E-38  -0.0000000E-38
      -0.0000000E-38  -0.0000000E-38  -0.0000000E-38   0.1000000E 01

TRANSFORMED MATRIX C(Z)

1 1    0.8000000E 01   0.2000000E 01   0.0000000E-38   0.0000000E-38   0.0000000E-38
1 2   -0.3000000E 01   0.2000000E 01   0.2000000E 01  -0.3000000E 01   0.1000000E 01
2 1   -0.3000000E 01   0.1000000E 01   0.0000000E-38   0.0000000E-38   0.0000000E-38
2 2    0.2100000E 03   0.5600000E 02   0.8400000E 02   0.8000000E 01   0.3000000E 01

ITERATION CONVERGED TO WITHIN 0.100E-07 IN    25 STEPS.  SOLUTION IS

1 1    0.7437588E 01   0.2000000E 01   0.0000000E-38   0.0000000E-38   0.0000000E-38
1 2   -0.3505871E 01   0.1000000E 01   0.0000000E-38   0.0000000E-38   0.0000000E-38
2 1   -0.3505871E 01   0.2047091E 00   0.2819858E 01  -0.3312866E 01   0.1000000E 01
2 2    0.1550601E 03   0.3550250E 02   0.8055831E 02   0.7166927E 01   0.3000000E 01

THE MATRIX FACTOR D(Z) IS

1 1    0.2712623E 01   0.7456289E 00   0.0000000E-38   0.0000000E-38   0.0000000E-38
1 2    0.0000000E-38   0.3713794E 00   0.1710987E 01  -0.1161541E 01   0.3936520E 00
2 1   -0.2815437E 00   0.8030637E-01   0.0000000E-38   0.0000000E-38   0.0000000E-38
2 2    0.1245231E 02   0.2851077E 01   0.6469345E 01   0.5755498E 00   0.2409191E 00

SOLUTION OF FACTORIZATION PROBLEM
G(S) = HT(-S)*H(S). H(S) IS

1 1    0.1729126E 01   0.9834969E 00   0.0000000E-38   0.0000000E-38   0.0000000E-38
1 2    0.8215482E-01   0.9320208E-01  -0.6625388E-01  -0.2900281E 00   0.1809251E 00
2 1   -0.1006187E 00  -0.1809251E 00   0.0000000E-38   0.0000000E-38   0.0000000E-38
2 2    0.1411825E 01   0.3337289E 01   0.3951294E 01   0.2768407E 01   0.9834969E 00

PRODUCT HT(-S)*H(S) IS

1 1    0.3000000E 01   0.0000000E-38  -0.1000000E 01  -0.0000000E-38  -0.0000000E-38
       0.0000000E-38   0.0000000E-38  -0.0000000E-38   0.0000000E-38
1 2   -0.0000000E-38  -0.3725290E-08   0.7450581E-08  -0.3725290E-08   0.1000000E 01
       0.0000000E-38   0.0000000E-38  -0.0000000E-38   0.0000000E-38
2 2    0.2000000E 01   0.0000000E-38  -0.2980232E-06  -0.5960464E-07  -0.1639128E-06
      -0.0000000E-38  -0.2682209E-06  -0.0000000E-38   0.9999999E 00
```

The answer obtained from Youla's exact solution by orthogonal transformation is

$$H(s) = \begin{bmatrix} .8814124165s & .472347490\,(1-s) & 0 \\ .472347490s & \begin{matrix}1.2356730345 \\ +\,.8814124165s\end{matrix} & 0 \\ 0 & -.5 & s \end{bmatrix}.$$

Thus, the greatest error in the factor is less than 10^{-7}.

3) Suppose

$$G(s) = \begin{bmatrix} 3-s^2 & s^4 \\ s^4 & 2+s^8 \end{bmatrix}$$

The computer output is shown in Fig. 3, and

$$H(s) = \begin{bmatrix} \begin{matrix}1.729126 \\ +.9834969s\end{matrix} & \begin{matrix}.08215482 + .09320208s \\ -.06625388s^2 \\ -\,.2900281s^3 \\ +\,.1809251s^4\end{matrix} \\ \begin{matrix}-.1006187 \\ -.1809251s\end{matrix} & \begin{matrix}1.411825 + 3.337289s \\ +\,3.951294s^2 \\ +\,2.768407s^3 \\ +\,.9834969s^4\end{matrix} \end{bmatrix}$$

The total execution time for these three examples was 1.8 seconds on the IBM 7094.

VI. Conclusion

An iterative algorithm for the spectral factorization of polynomial matrices has been derived and its convergence in regular cases established. The iteration is applied to a discrete-time polynomial obtained from the original continuous problem by a bilinear transformation. The result is then mapped back into the continuous domain.

The matrix recursive equation was obtained from a purely algebraic condition between a matrix polynomial and its factors. However, the equation turns out to be a special case of the Riccati difference equation employed in Kalman filtering.[12,14] Thus, known results on the Riccati equation were used to establish convergence of the algorithm.

The method seems to be superior to other techniques for factorization, both for the matrix and scalar case. One reason for this is that the roots of the polynomial, which are redundant information for the solution of the Wiener-Hopf equation, are never calculated. The factors are obtained directly. Thus, the algorithm produces only what is needed to solve the problem.

VII. Acknowledgments

The author is grateful for the help of T. C. Kelly, who programmed much of the algorithm and aided greatly in the checkout of the program. Thanks also to R. W. Koepcke for contributing the matrix inversion program.

References

1. N. Wiener, "Extrapolation, Interpolation, and Smoothing of Stationary Time Series," Wiley, New York, 1949.
2. Y. W. Lee, "Statistical Theory of Communication," Wiley, New York, 1960, pp. 355–396.
3. J. H. Laning and R. H Battin, "Random Processes in Automatic Control," McGraw-Hill, New York, 1956, pp. 269–290.
4. N. Wiener and P. Masani, "The Prediction Theory of Multivariate Stochastic Processes, II—The Linear Predictor," *Acta Mathematica* **99,** 93–97 (1958).
5. D. C. Youla, "On the Factorization of Rational Matrices" *IRE Transactions on Information Theory* **IT-7,** 172–189 (1961).
6. E. Wong and J. B. Thomas, "On the Multidimensional Prediction and Filtering Problem and the Factorization of Spectral Matrices," *Journal of the Franklin Institute* **272,** 87–99 (1961).
7. R. W. Brockett and M. D. Mesarović, "Synthesis of Linear Multivariable Systems," *Trans. AIEE, Applications and Industry* **62,** 216–221 (1962).
8. M. C. Davis, "Factoring the Spectral Matrix," *IEEE Transactions on Automatic Control* **AC-8,** 296–305 (1963).
9. P. Masani, "Recent Trends in Multivariate Prediction Theory," Report #637, January 1966, Mathematics Research Center, University of Wisconsin.
10. C. S. Lorens, "Factorization of Rational Matrices for Multiple Dimensional Stochastic Processes," *Preprints of the 1966 IFAC Conference,* London, England, pp. 8E.1–8E.7.
11. A. C. Riddle and B. D. Anderson, "Spectral Factorization—Computational Aspects," *IEEE Transactions on Automatic Control* **AC-11,** 764–765 (1966).
12. R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Trans. ASME, Series D, Journal of Basic Engineering* **82,** 35–45 (1960).
13. T. Nishimura, "On the a priori Information in Sequential Estimation Problems," *IEEE Transactions on Automatic Control* **AC-11,** 197–204 (1966).
14. W. Tuel, "Canonical Forms for Linear Systems-II; Application to Optimal Filtering and Control," IBM Research Report RJ-390, September 1966.
15. B. D. Anderson, "An Algebraic Solution to the Spectral Factorization Problem," *IEEE Trans. Automatic Control* **AC-12,** 410–414 (1967).
16. F. Csáki and P. Fischer, "On the Spectrum Factorization," *Acta Technica Academiae Scientiarium Hungaricae* **58,** 145–168 (1967).

Received March 6, 1967.