The Inverse of a Matrix Polynomial

C. E. Langenhop
Department of Mathematics
Southern Illinois University
Carbondale, Illinois 62901

Submitted by Hans Schwerdtfeger

ABSTRACT

An explicit representation is obtained for $P(z)^{-1}$ when $P(z)$ is a complex $n \times n$ matrix polynomial in z whose coefficient of the highest power of z is the identity matrix. The representation is a sum of terms involving negative powers of $z-\lambda$ for each λ such that $P(\lambda)$ is singular. The coefficients of these terms are generated by sequences u_k, v_k of $1 \times n$ and $n \times 1$ vectors, respectively, which satisfy $u_1 \neq 0$, $v_1 \neq 0$, $\sum_{h=0}^{k-1}(1/h!)u_{k-h}P^{(h)}(\lambda)=0$, $\sum_{h=0}^{k-1}(1/h!)P^{(h)}(\lambda)v_{k-h}=0$, and certain orthogonality relations. In more general cases, including that when $P(z)$ is analytic at λ but not necessarily a polynomial, the terms in the representation involving negative powers of $z-\lambda$ provide the principal part of the Laurent expansion for $P(z)^{-1}$ in a punctured neighborhood of $z=\lambda$.

1. INTRODUCTION

In [7] and [8] we have considered the inverse of $A+zB$ near $z=0$ when A and B are $n \times n$ matrices over the complex numbers and A is singular. In [6] the inverse of the matrix polynomial Az^2+Bz+C was considered using a variation of Leverrier's algorithm [4, pp. 87–88]. In [5] C. J. Hegedüs has studied the problem of finding the inverse of an $n \times n$ matrix $A(z)$ the elements of which are analytic functions of a complex variable. In some applications referred to there, the main interest resides in determining the residue at the poles (or other isolated singularities) of $A(z)^{-1}$. In other applications, especially when the elements of $A(z)$ are rational [2, 3], one is interested in complete knowledge of $A(z)^{-1}$. Although the theory developed below can lead to a complete representation of $A(z)^{-1}$ in such cases, it is perhaps of greater significance in regard to residue evaluation or the determination of the principal part of the Laurent expansion of $A(z)^{-1}$ in a punctured neighborhood of a pole of this function.

By translation of the variable z we may suppose that $z=0$ is the point at which we disire the residue of $A(z)^{-1}$. Thus we assume

$$A(z)=A_0+A_1z+A_2z^2+\cdots, \tag{1.1}$$

where the series converges on the disk $|z|<r$ for some $r>0$. The point $z=0$ is an isolated singularity of $A(z)^{-1}$ if $\det A(0)=0$ but $\det A(z)\not\equiv 0$. The singularity must then be a pole, since $\operatorname{adj} A(z)$ and $\det A(z)$ are analytic at 0 with $z=0$ a zero of $\det A(z)$ of finite order. Suppose for $|z|<r$ that

$$\det A(z)=z^{\mu}(d_{\mu}+d_{\mu+1}z+\cdots), \tag{1.2}$$

where $\mu\geqslant 1$ and $d_{\mu}\neq 0$. Then the elements of $A(z)^{-1}$ can have at most a pole of order μ at $z=0$.

For any m let us write

$$A(z)=A_{(m)}(z)+R_{(m)}(z), \tag{1.3}$$

where

$$A_{(m)}(z)=A_0+A_1z+\cdots+A_mz^m \tag{1.4}$$

and

$$R_{(m)}(z)=A_{m+1}z^{m+1}+\cdots \tag{1.5}$$

for $|z|<r$. It is not difficult to see that

$$\det A(z)=\det A_{(m)}(z)+z^{m+1}\Delta(z), \tag{1.6}$$

where $\Delta(z)$ is a scalar function analytic at $z=0$. Indeed, $\det A_{(m)}(z)$ is obtained by truncating all the elements in $A(z)$ after the terms in z^m, so $\det A_{(m)}(z)$ must agree with $\det A(z)$ through terms of order z^m. From (1.6) it is clear that if $m\geqslant\mu$, then $z=0$ is a zero of order μ of $\det A_{(m)}(z)$. On a punctured neighborhood of 0 we may thus write, provided $m\geqslant\mu$,

$$A(z)^{-1}=A_{(m)}(z)^{-1}[I-B(z)]^{-1} \tag{1.7}$$

with

$$B(z) = -R_{(m)}(z)A_{(m)}(z)^{-1}. \tag{1.8}$$

Assuming $m \geqslant \mu$, we see that the elements of $A_m(z)^{-1}$ have at most a pole of order μ at 0, so $B(z) \to 0$ as $z \to 0$; in effect, the elements of $B(z)$ are analytic with a zero of order at least $m+1-\mu$ at $z=0$. Hence, to compute the principal part of $A(z)^{-1}$ at $z=0$ it suffices to take $m \geqslant 2\mu-1$ and compute the same for $A_{(m)}(z)^{-1}$. Indeed,

$$A(z)^{-1} = A_{(m)}(z)^{-1} + \sum_{k=1}^{\infty} A_{(m)}(z)^{-1}B(z)^k \tag{1.9}$$

near $z=0$, and for $m \geqslant 2\mu-1$ the term represented by the summation on the right is, in effect, analytic at $z=0$. Hence, if

$$A_{(m)}(z)^{-1} = \sum_{k=1}^{\mu} G_k z^{-k} + G_{(m)}(z), \tag{1.10}$$

where the G_k are constant $n \times n$ matrices and $G_{(m)}(z)$ is analytic at 0, then

$$A(z)^{-1} = \sum_{k=1}^{\mu} G_k z^{-k} + G(z), \tag{1.11}$$

where $G(z)$ is analytic at $z=0$.

The above argument justifies the following formal statement:

Proposition 1.1. *Let $A(z)$ and $\tilde{A}(z)$ be $n \times n$ matrix functions of z which are analytic at $z=0$ and satisfy*

$$A(z) - \tilde{A}(z) = (z-\lambda)^d K(z), \tag{1.12}$$

where $K(z)$ is analytic at $z=\lambda$. If $\det \tilde{A}(z)$ has a zero of order $\mu \geqslant 1$ at $z=\lambda$ and $d \geqslant 2\mu$, then the principal parts at $z=\lambda$ of $A(z)^{-1}$ and $\tilde{A}(z)^{-1}$ are identical; that is, $A(z)^{-1} - \tilde{A}(z)^{-1}$ can be extended to an analytic function at $z=\lambda$.

From the above discussion we see that to obtain the principal part of $A(z)^{-1}$ at $z=0$ it suffices to obtain it in the case when $A(z)$ is a polynomial, say $A(z)=P(z)$, where

$$P(z)=P_0+P_1z+\cdots+P_mz^m$$

and $\det P(z)\not\equiv 0$. Hence we may assume $P_m\neq 0$, and by virture of Proposition 1.1, we may for our purposes assume $P_m=I_n$, the $n\times n$ identity matrix. One method for computing $P(z)^{-1}$ is to reduce $P(z)$ to diagonal form by pre- and post-multiplication by elementary polynomial matrices with constant non-zero determinants [4, Chapter VI]. The inverses of such matrices are likewise polynomials in z and determination of the principal part at $z=0$ of the inverse of the resulting diagonal matrix may be carried out by the usual scalar techniques. Below we develop other theoretical relations which, conceivably, could be more easily applied in some situations.

2. INVERSE OF A MATRIX POLYNOMIAL

We consider here $n\times n$ matrix polynomials

$$P(z)=P_0+P_1z+\cdots+P_mz^m, \tag{2.1}$$

in which the P_k are $n\times n$ over the complex numbers $\mathbf{C}$ and

$$P_m=I_n, \quad n\times n \text{ identity}. \tag{2.2}$$

Associated with $P(z)$ we introduce a generalized companion matrix of size $\delta\times\delta$, where $\delta=nm$; in block form we define

$$\mathcal{P}=\begin{bmatrix} 0 & I_n & 0 & \cdots & 0 \\ 0 & 0 & I_n & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & I_n \\ -P_0 & -P_1 & -P_2 & \cdots & -P_{m-1} \end{bmatrix}. \tag{2.3}$$

Then one may show that

$$\det P(z)=\det(zI_\delta-\mathcal{P}). \tag{2.4}$$

This also follows from (2.17) below using a Laplace expansion by $n\times n$ minors in the first n columns.

Our aim is to relate $P(z)^{-1}$ to $(zI_\delta - \mathcal{P})^{-1}$. In doing this the following terminology is convenient.

DEFINITION 2.1. The complex number λ is a singular value for a matrix polynomial $P(z)$ if $\det P(\lambda)=0$; an $n\times 1$ $[1\times n]$ vector v over $\mathbf{C}$ is a singular column [row] vector of $P(z)$ belonging to λ if $v\neq 0$ and $P(\lambda)v=0$ $[vP(\lambda)=0]$.

From (2.4) it is evident that λ is a singular value for $P(z)$ if and only if λ is an eigenvalue of $\mathcal{P}$. As we shall see, one can generate eigenvectors of $\mathcal{P}$ from the singular vectors of $P(z)$.

Our representation for $P(z)^{-1}$ will be derived from a representation of $(zI_\delta - \mathcal{P})^{-1}$ obtained through a Jordan form for $\mathcal{P}$ using a decomposition of the type described in [1]. Let us first introduce certain $\delta\times n$ matrices in block form as follows:

$$E_k = [0,\dots,0,I_n,0,\dots,0]^T, \qquad k=1,\dots,m, \tag{2.5}$$

in which the zeros denote $n\times n$ zero matrices and I_n appears as the kth block. For convenience we also introduce two $\delta\times n$ matrices of zeros denoted by E_0 and E_{m+1}:

$$E_0=0 \quad (\delta\times n), \qquad E_{m+1}=0 \quad (\delta\times n). \tag{2.6}$$

Now define

$$F(z)=E_1+zE_2+\cdots+z^{m-1}E_m. \tag{2.7}$$

From (2.3), (2.5), and (2.6) we readily get

$$\mathcal{P}E_k=E_{k-1}-E_mP_{k-1}, \qquad k=1,\dots,m. \tag{2.8}$$

Using this and the fact that $P_m=I_n$, we may verify the identity [cf. (2.1) and (2.7)]

$$\mathcal{P}F(z)=zF(z)-E_mP(z). \tag{2.9}$$

Since $E_m^TE_m=I_n$, (2.9) implies the following representation for $P(z)$:

$$P(z)=E_m^T(zI_\delta-\mathcal{P})F(z). \tag{2.10}$$

To obtain a representation for $P(z)^{-1}$ we use (2.9) and the results obtained by differentiating it $m-1$ times. One readily finds for $k \geqslant 0$

$$\mathcal{P} F^{(k)}(z) = zF^{(k)}(z) + kF^{(k-1)}(z) - E_m P^{(k)}(z). \tag{2.11}$$

Now introduce the $\delta \times \delta$ matrix polynomial

$$\mathcal{F}(z) = \left[F(z), \quad F'(z), \quad \frac{1}{2!} F''(z), \ldots, \frac{1}{(m-1)!} F^{(m-1)}(z) \right]. \tag{2.12}$$

When $\mathcal{F}(z)$ is partitioned into $n \times n$ blocks, we have

$$i,j \quad \text{block in} \quad \mathcal{F}(z) = \binom{i-1}{j-1} z^{i-j} I_n \tag{2.13}$$

with the usual conventions

$$\binom{k}{0} = 1, \quad k = 0, 1, \ldots; \qquad \binom{k}{j} = 0 \quad \text{for} \quad k < j. \tag{2.14}$$

Note that $\mathcal{F}(z)$ is lower triangular with ones on the main diagonal, so $\mathcal{F}(z)^{-1}$ exists for all z. Using (2.11) and (2.12), one may verify that

$$(zI_\delta - \mathcal{P})\mathcal{F}(z) = E_m \left[P(z), \quad P'(z), \ldots, \frac{1}{(m-1)!} P^{(m-1)}(z) \right] - \mathcal{F}(z)\mathcal{N} \tag{2.15}$$

where $\mathcal{N}$ is the $\delta \times \delta$ nilpotent matrix

$$\mathcal{N} = [E_0, E_1, \ldots, E_{m-1}]. \tag{2.16}$$

From (2.13) we see that $E_m = \mathcal{F}(z) E_m$. Using this in (2.15), we find

$$\mathcal{F}(z)^{-1}(zI_\delta - \mathcal{P})\mathcal{F}(z) = \begin{bmatrix} 0 & -I_n & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -I_n \\ P(z) & P'(z) & \cdots & \frac{1}{(m-1)!} P^{(m-1)}(z) \end{bmatrix}. \tag{2.17}$$

It is not difficult to find the inverse of the matrix on the right in (2.17).

Indeed, when z is not a singular value for $P(z)$ we have

$$\mathcal{F}(z)^{-1}(zI_\delta - \mathcal{P})^{-1}\mathcal{F}(z)$$

$$=\begin{bmatrix} P(z)^{-1}P'(z) & \cdots & P(z)^{-1}\frac{1}{(m-1)!}P^{(m-1)}(z) & P(z)^{-1} \\ -I_n & \cdots & 0 & 0 \\ \vdots & & \vdots & \\ 0 & \cdots & -I_n & 0 \end{bmatrix}. \tag{2.18}$$

It follows that

$$P(z)^{-1} = E_1^T\mathcal{F}(z)^{-1}(zI_\delta - \mathcal{P})^{-1}\mathcal{F}(z)E_m. \tag{2.19}$$

A simpler presentation is possible.

THEOREM 2.1. *If $P_m = I_n$ and z is not a singular value for $P(z)$, then*

$$P(z)^{-1} = E_1^T(zI_\delta - \mathcal{P})^{-1}E_m. \tag{2.20}$$

Proof. We have already used the relation $E_m = \mathcal{F}(z)E_m$, so it remains to show that $E_1^T\mathcal{F}(z)^{-1} = E_1^T$. To this end we introduce $\mathcal{B} = [E_2, \dots, jE_{j+1}, \dots, mE_{m+1}]$, in which the last n columns are zero by (2.6). From (2.12) it readily follows that $\mathcal{F}'(z) = \mathcal{F}(z)\mathcal{B}$, since $F^{(m)}(z) \equiv 0$. Moreover, $\mathcal{F}(0) = I_\delta$ by virtue of (2.13), so $\mathcal{F}(z) = e^{\mathcal{B}z}$. Hence $\mathcal{F}(z)^{-1} = \mathcal{F}(-z)$ and $E_1^T\mathcal{F}(z)^{-1} = E_1^T\mathcal{F}(-z) = E_1^T$ by (2.13). Thus (2.19) reduces to (2.20). ■

3. REPRESENTATION FOR $P(z)^{-1}$

Now let $\mathcal{V}$ be a nonsingular $\delta \times \delta$ matrix such that

$$\mathcal{V}^{-1}\mathcal{P}\mathcal{V} = \mathcal{J} = \operatorname{diag}(J_1, \dots, J_r) \tag{3.1}$$

is a Jordan form. That is, we suppose J_k is $\sigma_k \times \sigma_k$, $\sigma_k \geqslant 1$, $\sigma_1 + \cdots + \sigma_r = \delta$, and

$$J_k = [\lambda_k] \quad \text{or} \quad J_k = \lambda_k I + N, \tag{3.2}$$

where I is an identity matrix and N the same size nilpotent matrix of the form

$$N=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{3.3}$$

Let

$$\mathcal{U}=\mathcal{V}^{-1}. \tag{3.4}$$

We partition $\mathcal{U}$ and $\mathcal{V}$ as follows:

$$\mathcal{U}=\begin{bmatrix} U_{11} & \cdots & U_{1m} \\ \vdots & & \vdots \\ U_{r1} & \cdots & U_{rm} \end{bmatrix}, \quad \mathcal{V}=\begin{bmatrix} V_{11} & \cdots & V_{1r} \\ \vdots & & \vdots \\ V_{m1} & \cdots & V_{mr} \end{bmatrix}, \tag{3.5}$$

where U_{ij} is $\sigma_i \times n$ and V_{ij} is $n \times \sigma_j$. Since $\mathcal{P}\mathcal{V}=\mathcal{V}\mathcal{J}$, it follows that for $j=1,\dots,r$

$$V_{i+1,j}=V_{ij}J_j, \qquad i=1,\dots,m-1 \tag{3.6}$$

and

$$P_0V_{ij}+\cdots+P_{m-1}V_{mj}+V_{mj}J_j=0. \tag{3.7}$$

From (3.6) we find

$$V_{ij}=V_{1j}J_j^{i-1}, \qquad i=2,\dots,m, \tag{3.8}$$

and these substituted in (3.7) give

$$P_0V_{1j}+P_1V_{1j}J_j+\cdots+P_mV_{1j}J_j^m=0, \qquad j=1,\dots,r, \tag{3.9}$$

since $P_m=I_n$ by (2.2). Similarly, since $\mathcal{U}\mathcal{P}=\mathcal{J}\mathcal{U}$, we get for $i=1,\dots,r$

$$U_{i,j-1}=U_{im}P_{j-1}+J_iU_{ij}, \qquad j=2,\dots,m, \tag{3.10}$$

and

$$U_{im}P_0 + J_i U_{i1} = 0. \tag{3.11}$$

Equations (3.10) may be used to express U_{ij} in terms of U_{im}, $j = 1,\dots,m-1$. Substituting the result for U_{i1} into (3.11), we get, since $P_m = I_n$,

$$U_{im}P_0 + J_i U_{im}P_1 + \cdots + J_i^m U_{im}P_m = 0, \qquad i = 1,\dots,r.$$

Turning to our representation (2.20), we have by virtue of (3.1)

$$P(z)^{-1} = E_1^T \mathcal{V}(zI_\delta - \mathcal{J})^{-1}\mathcal{V}^{-1}E_m. \tag{3.12}$$

But

$$(zI_\delta - \mathcal{J})^{-1} = \operatorname{diag}\big((zI - J_1)^{-1},\dots,(zI - J_r)^{-1}\big),$$

where the identities on the right are the same size as the associated J_k. From (3.5) and (3.4) we have

$$E_1^T\mathcal{V} = [\,V_{11},\dots,V_{1r}\,], \qquad \mathcal{V}^{-1}E_m = \begin{bmatrix} U_{1m} \\ \vdots \\ U_{rm} \end{bmatrix},$$

so substituting into (3.12) we find

$$P(z)^{-1} = \sum_{k=1}^{r} V_{1k}(zI - J_k)^{-1}U_{km} \tag{3.13}$$

for $z \neq \lambda_i$, $i = 1,\dots,r$.

Our ultimate aim is to describe the structure of the matrices V_{1k} and U_{km} in (3.13). Before doing this, however, let us state formally the way these are related to $\mathcal{V}$ and $\mathcal{U} = \mathcal{V}^{-1}$. In this connection the following terminology is useful.

DEFINITION 3.1. Let V be $\delta \times \sigma$ and U be $\sigma \times \delta$, and let $J = \lambda I + N$ [N as in (3.3)] be a $\sigma \times \sigma$ Jordan block with λ an eigenvalue of $\mathcal{P}$. We say V is a right eigenmatrix of $\mathcal{P}$ belonging to the eigenvalue λ if the first column of V is nonzero and

$$\mathcal{P}V = VJ. \tag{3.14}$$

We say U is a left eigenmatrix of $\mathcal{P}$ belonging to λ if the last row of U is nonzero and

$$U\mathcal{P} = JU.$$

Note that V is a right eigenmatrix for $\mathcal{P}$ if and only if the columns of V in order form a Jordan chain for $\mathcal{P}$. (See [4], p. 165.) Similarly U is a left eigenmatrix for $\mathcal{P}$ if and only if the rows of U in *reverse* order form a Jordan chain of row vectors for $\mathcal{P}$.

Some of the relations derived earlier are contained in the following:

THEOREM 3.1. *Let $P_m = I_n$. Let V be $\delta\times\sigma$ and U be $\sigma\times\delta$, and let $J=\lambda I+N$ be a $\sigma\times\sigma$ Jordan block. Define*

$$V_i = E_i^T V, \qquad U_i = UE_i, \qquad i=1,\dots,m. \tag{3.15}$$

Then V is a right eigenmatrix of $\mathcal{P}$ belonging to λ if and only if

$$V_i = V_1 J^{i-1}, \qquad i=2,\dots,m, \tag{3.16}$$

where

$$P_0V_1 + P_1V_1J + \cdots + P_mV_1J^m = 0, \tag{3.17}$$

and the first column of V_1 is nonzero. Also U is a left eigenmatrix of $\mathcal{P}$ belonging to λ if and only if

$$U_i = U_mP_i + JU_{i+1}, \qquad i=1,\dots,m-1, \tag{3.18}$$

where

$$U_mP_0 + JU_mP_1 + \cdots + J^mU_mP_m = 0 \tag{3.19}$$

and the last row of U_m is nonzero.

Proof. The derivations of (3.8) and (3.9) given above are essentially the proof that (3.16) and (3.17) are implied by (3.14). Now from (3.16) one sees that the first column of V_i is λ^{i-1} times the first column of V_1, $i=2,\dots,m$. Hence the first column of V is nonzero if and only if the first column of V_1 is nonzero. Now suppose (3.16) and (3.17) hold. Then we have

$$V_{i+1} = V_iJ, \qquad i=1,\dots,m-1$$

and

$$P_0V_1 + P_1V_2 + \cdots + P_{m-1}V_m + V_mJ = 0,$$

since $P_m = I_n$. But these last two relations are equivalent to (3.14). This completes the proof of the assertion in the theorem regarding V. That regarding U is likewise straightforward, so we omit the details except to note that (3.18) is equivalent to

$$U_i = U_mP_i + JU_mP_{i+1} + \cdots + J^{m-i}U_mP_m, \tag{3.20}$$

$i = 1,\dots,m-1$. ■

4. DEVELOPMENT

In this section we consider the representation (3.13) for $P(z)^{-1}$ in finer detail. It will be convenient at first to suppress the index k in considering the terms $V_{1k}(zI-J_k)^{-1}U_{km}$ in the sum yielding $P(z)^{-1}$. Accordingly, we suppose $J = \lambda I + N$ is a $\sigma\times\sigma$ Jordan block [cf. (3.3)]. Then one has

$$(zI-J)^{-1} = \begin{bmatrix} \frac{1}{z-\lambda} & \frac{1}{(z-\lambda)^2} & \cdots & \frac{1}{(z-\lambda)^\sigma} \\ 0 & \frac{1}{z-\lambda} & \cdots & \frac{1}{(z-\lambda)^{\sigma-1}} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{1}{z-\lambda} \end{bmatrix}. \tag{4.1}$$

Now let V_1 be $n\times\sigma$ and U_m be $\sigma\times n$, and let their columns and rows, respectively, be indexed as in

$$V_1 = [v_1, v_2, \dots, v_\sigma], \qquad U_m = \begin{bmatrix} u_\sigma \\ u_{\sigma-1} \\ \vdots \\ u_1 \end{bmatrix}. \tag{4.2}$$

Then we may write

$$V_1(zI-J)^{-1}U_m = \sum_{j=1}^{\sigma} \frac{1}{(z-\lambda)^j} W_j, \tag{4.3}$$

where

$$W_j = \sum_{h=1}^{\sigma+1-j} v_h u_{\sigma+2-h-j}, \qquad j=1,\dots,\sigma. \tag{4.4}$$

In (4.4) the $n\times n$ coefficient matrices in the expansion (4.3) are given in terms of the columns and rows of V_1 and U_m, respectively.These in turn satisfy certain relations defined in terms of $P(z)$ and its derivatives at $z=\lambda$. Specifically we have the following result.

THEOREM 4.1. *Let V_1 be $n\times\sigma$ and U_m be $\sigma\times m$, and let $J=\lambda I+N$ be a $\sigma\times\sigma$ Jordan block. Let the columns and rows of V_1 and U_m, respectively, be indexed as in* (4.2). *Then V_1 satisfies* (3.17) *if and only if*

$$\sum_{h=0}^{k-1} \frac{1}{h!} P^{(h)}(\lambda) v_{k-h} = 0, \qquad k=1,\dots,\sigma. \tag{4.5}$$

Also U_m satisfies (3.19) *if and only if*

$$\sum_{h=0}^{k-1} \frac{1}{h!} u_{k-h} P^{(h)}(\lambda) = 0, \qquad k=1,\dots,\sigma. \tag{4.6}$$

Proof. From (4.2) and (3.3) one finds that the kth column of $V_1 J^j$ is given by

$$\sum_{h=0}^{k-1} \binom{j}{h} \lambda^{j-h} v_{k-h}.$$

Hence (3.17) holds if and only if

$$\sum_{j=0}^{m} P_j \sum_{h=0}^{k-1} \binom{j}{h} \lambda^{j-h} v_{k-h} = \sum_{h=0}^{k-1} \frac{1}{h!} P^{(h)}(\lambda) v_{k-h} = 0$$

for $k=1,\dots,\sigma$. This is (4.5). The proof of the equivalence of (3.19) and (4.6) is similar. ■

The above result suggests that some terminology be introduced for sequences such as $v_1,\dots,v_\sigma$ or $u_1,\dots,u_\sigma$. We propose the following.

DEFINITION 4.1. The $n\times 1$ vectors $v_1,\dots,v_\sigma$ [or $1\times n$ vectors $u_1,\dots,u_\sigma$] form a singular column [row] sequence for $P(z)$ belonging to the singular value λ if $v_1\neq 0$ and (4.5) holds [$u_1\neq 0$ and (4.6)) holds].

Note that by this definition v_1 (u_1) is a singular column (row) vector of $P(z)$ belonging to λ as in Definition 2.1. The same singular value may be represented in two or more of the Jordan blocks J_k in (3.1). In any case, the right and left eigenmatrices defined through (4.2), (3.16) and (3.18) by the corresponding singular column and row sequences for $P(z)$ must satisfy

$$\mathfrak{U}\,\mathfrak{V}=I_\delta. \tag{4.7}$$

We will refer to the relations among singular column and row sequences which are implicit in (4.7) as the *normalizing conditions*.

Using the partitioning in (3.7) we define

$$U^{(k)}=\left[\,U_{k1},\dots,U_{km}\right],\qquad V^{(k)}=\begin{bmatrix} V_{1k}\\ \vdots \\ V_{mk}\end{bmatrix}. \tag{4.8}$$

Thus $U^{(k)}$ is $\sigma_k\times\delta$, $V^{(k)}$ is $\delta\times\sigma_k$, and these are left and right eigenmatrices of $\mathcal{P}$, respectively, belonging to the singular value λ_k of $P(z)$. In terms of them, Eq. (4.7) is equivalent to

$$U^{(k)}V^{(k)}=I_{\sigma_k},\qquad k=1,\dots,r, \tag{4.9}$$

and

$$U^{(k')}V^{(k)}=0,\qquad k'\neq k,\quad k,k'=1,\dots,r. \tag{4.10}$$

If $\lambda_{k'}\neq\lambda_k$ for some k,k' with $k'\neq k$, then the corresponding condition (4.10) will hold automatically. Specifically we have, dropping the indices k, k':

THEOREM 4.2. *Let the $\delta\times\sigma$ matrix V be a right eigenmatrix for $\mathcal{P}$ belonging to the eigenvalue λ associated with the $\sigma\times\sigma$ Jordan block $J=\lambda I+N$ [cf. (3.3)]. Let the $\sigma'\times\delta$ matrix U be a left eigenmatrix of $\mathcal{P}$ belonging to the eigenvalue λ' associated with the $\sigma'\times\sigma'$ Jordan block $J'=\lambda'I'+N'$. If $\lambda'\neq\lambda$, then $UV=0$.*

Proof. From $\mathcal{P}V=VJ$ and $U\mathcal{P}=J'U$ it readily follows that $UVJ=J'UV$. This in turn implies

$$(\lambda-\lambda')UV=N'UV-UVN. \tag{4.11}$$

From (4.11) it is not difficult, by examining the element by element equations in proper sequence, to show that $UV=0$ if $\lambda-\lambda'\neq0$. ∎

In somewhat the same manner one finds that when $\lambda_{k'} = \lambda_k$ for some k, k' with $k' \neq k$, the relation (4.10) is equivalent to fewer conditions than the number of elements in the product $U^{(k')}V^{(k)}$; a similar remark applies also to (4.9). Indeed, dropping subscripts and using the basic assumptions given in Theorem 4.2, we see that if $\lambda' = \lambda$ then by (4.11) we must have

$$N'UV = UVN. \tag{4.12}$$

Let us use x_{ij} to denote the element in the ith row and jth column of UV. Then (4.12) is equivalent to

$$x_{ij} = x_{i-1,j-1}, \qquad i = 2,\dots,\sigma', \quad j = 2,\dots,\sigma, \tag{4.13}$$

$$x_{\sigma' j} = 0, \qquad j = 1,\dots,\sigma - 1 \tag{4.14}$$

and

$$x_{i1} = 0, \qquad i = 2,\dots,\sigma'. \tag{4.15}$$

From these it follows that

$$x_{ij} = 0, \qquad i - j > \min\{0, \sigma' - \sigma\}, \tag{4.16}$$

$$x_{ij} = c_{i-j}, \qquad i - j \leqslant \min\{0, \sigma' - \sigma\}. \tag{4.17}$$

Accordingly we have the following result.

THEOREM 4.3. *Let V, U satisfy the basic assumptions given in Theorem 4.2, and let $UV = (x_{ij})$. If $\lambda' = \lambda$, then $UV = 0$ if and only if*

$$x_{i\sigma} = 0, \qquad 1 \leqslant i \leqslant \sigma + \min\{0, \sigma' - \sigma\}. \tag{4.18}$$

If $\lambda' = \lambda$ and $\sigma' = \sigma$, then $UV = I_\sigma$ if and only if

$$x_{i\sigma} = 0, \qquad i = 1,\dots,\sigma - 1, \tag{4.19}$$

and

$$x_{\sigma\sigma} = 1. \tag{4.20}$$

We now wish to represent the elements of the various products $U^{(k')}V^{(k)}$ in terms of the respective singular row and column sequences. From Theorem 4.2 it is evident that we need concern ourselves only with the cases for which $\lambda_{k'}=\lambda_k$. Accordingly we again suppress the indices k, k', but since the same eigenvalue λ of $\mathcal{P}$ may appear in different sized blocks in $\mathcal{J}$ [cf. (3.1)], we assume V is $\delta \times \sigma$ and U is $\sigma' \times \delta$ as in Theorem 4.2. Using the notation in (3.15), we have

$$UV=\sum_{k=1}^{m} U_k V_k, \tag{4.21}$$

and by (3.16) and (3.20) this becomes

$$UV=\sum_{k=1}^{m}\sum_{h=k}^{m}(J')^{h-k}U_m P_h V_1 J^{k-1}. \tag{4.22}$$

By (4.2) and the relation $J'=\lambda I'+N$ (recall we are assuming $\lambda'=\lambda$), the ith row of $(J')^{h-k}U_m$ is

$$\sum_{s=0}^{h-k}\binom{h-k}{s}\lambda^{h-k-s}u_{\sigma'+1-i-s}$$

if we adopt the convention

$$u_i=0 \qquad \text{if} \quad i\leqslant 0. \tag{4.23}$$

Similarly the jth column of $V_1 J^{k-1}$ is

$$\sum_{p=0}^{j-1}\binom{k-1}{p}\lambda^{k-1-p}v_{j-p}.$$

Hence

$$x_{ij}=\sum_{k=1}^{m}\sum_{h=k}^{m}\sum_{s=0}^{h-k}\sum_{p=0}^{j-1}\binom{h-k}{s}\binom{k-1}{p}\lambda^{h-s-p-1}u_{\sigma'+1-i-s}P_h v_{j-p}.$$

Interchanging the order of summation on the right, we may write this as

$$x_{ij}=\sum_{s=0}^{m-1}\sum_{p=0}^{j-1}u_{\sigma'+1-i-s}\sum_{h=s+1}^{m}\sum_{k=1}^{h-s}\binom{h-k}{s}\binom{k-1}{p}\lambda^{h-s-p-1}P_h v_{j-p}.$$

Use of an identity between binomial coefficients leads to

$$x_{ij}=\sum_{s=0}^{m-1}\sum_{p=1}^{j}u_{\sigma'+1-i-s}\frac{1}{(s+p)!}P^{(s+p)}(\lambda)v_{j+1-p}. \tag{4.24}$$

We are now in a position to specify the normalizing conditions equivalent to (4.7). For notational convenience let us introduce a sequence of $n\times 1$ column vectors

$$w_s=\sum_{p=1}^{\sigma}\frac{1}{(p+s)!}P^{(p+s)}(\lambda)v_{\sigma+1-p} \tag{4.25}$$

associated with the singular column sequence $v_1,\dots,v_\sigma$ of $P(z)$ belonging to the singular value λ of $P(z)$. Then by (4.24) we may write

$$x_{\sigma'+1-i,\sigma}=\sum_{s=0}^{m-1}u_{i-s}w_s. \tag{4.26}$$

DEFINITION 4.2. Let $P(z)$ be any $n\times n$ polynomial as in (2.1), and let λ be a singular value of $P(z)$ with multiplicity μ as a zero of $\det P(z)$. Let $\sigma(\alpha)$, $\alpha=1,\dots,\tau$, be a set of positive integers such that

$$\sigma(1)+\cdots+\sigma(\tau)=\mu \tag{4.27}$$

and let $u_i(\alpha)$, $v_i(\alpha)$, $i=1,\dots,\sigma(\alpha)$ be, respectively, singular row and column sequences of $P(z)$ belonging to λ which satisfy

$$\sum_{s=0}^{m-1}u_{i-s}(\alpha')w_s(\alpha)=\begin{cases}1 & \text{if } \alpha'=\alpha,\quad i=1,\\ 0 & \text{if } \alpha'=\alpha,\quad 2\leqslant i\leqslant\sigma(\alpha),\\ 0 & \text{if } \alpha'\neq\alpha,\quad 1+\max\{0,\sigma(\alpha')-\sigma(\alpha)\}\leqslant i\leqslant\sigma(\alpha'),\end{cases} \tag{4.28}$$

where $w_s(\alpha)$ is defined in terms of the $v_i(\alpha)$ by (4.25) with $\sigma=\sigma(\alpha)$, and where $u_i(\alpha)=0$ for $i\leqslant 0$. Then we call the sequences $u_i(\alpha)$, $v_i(\alpha)$, $i=1,\dots,\sigma(\alpha)$, $\alpha=1,\dots,\tau$, a complete normalized set of singular row and column sequences of $P(z)$ belonging to λ.

It should be clear from our earlier developments that when $P_m=I_n$, the sizes of the Jordan blocks J_k [in (3.1)] for which $\lambda_k=\lambda$ are positive integers σ_k the sum of which gives the multiplicity of λ as a zero of $\det P(z)$. To each such block we have corresponding left and right eigenmatrices $U^{(k)}$, $V^{(k)}$ of $\mathcal{P}$ as in (4.8) satisfying the hypotheses of Theorem 4.3. The resulting normalizing conditions (4.18)–(4.20) translate into (4.28) by virtue of (4.26). The various Jordan blocks for which $\lambda_k=\lambda$ are indexed by α in Definition 4.2. We may summarize these remarks as follows.

THEOREM 4.4. *Let the $n\times n$ polynomial $P(z)$ given by* (2.1) *have $P_m = I_n$. Let the distinct singular values for $P(z)$ be $\lambda_1,\dots,\lambda_t$ with multiplicities $\mu_1,\dots,\mu_t$ as zeros of $\det P(z)$:*

$$\mu_1+\cdots+\mu_t=\delta=nm.$$

Then for each λ_k, $k=1,\dots,t$, there is a complete normalized set $u_i(\lambda_k,\alpha)$, $v_i(\lambda_k,\alpha)$, $i=1,\dots,\sigma_k(\alpha)$, $\alpha=1,\dots,\tau_k$, of singular row and column sequences of $P(z)$ belonging to λ_k. Moreover, any such collection generates a family of left and right eigenmatrices of $\mathcal{P}$ via Eqs. (4.2), (4.8), (3.16) *and* (3.18) *with $J=\lambda_k I+N$ of size $\sigma_k(\alpha)$. When these are combined as in* (3.5), *they generate $\delta\times\delta$ matrices $\mathcal{U}$ and $\mathcal{V}$ satisfying* (4.7) *and* (3.1).

Proof. Aside from our earlier discussion we need only point out that by Theorem 4.2 the cases $\lambda_{k'}\neq\lambda_k$ contribute no normalizing conditions. In relation to (3.1) we point out that $\tau_1+\cdots+\tau_t=r$, inasmuch as the Jordan blocks there do not necessarily correspond to distinct eigenvalues of $\mathcal{P}$. ∎

5. FINAL RESULT

We can now describe $P(z)^{-1}$ precisely in terms of complete normalized singular row and column sequences of $P(z)$ when $P_m=I_n$. When P_m is nonsingular we can apply our previous results to $\hat{P}(z)=P_m^{-1}P(z)$. One gets the same form then for $P(z)^{-1}$ in terms of a complete normalized set of singular row and column sequences of $P(z)$. When P_m is singular we may appeal to Proposition 1.1 and applly our result to $\tilde{P}(z)=P(z)+I_n(z-\lambda)^d$ for sufficiently large d in order to obtain a representation for the principal part of $P(z)^{-1}$ at a singular value λ of $P(z)$. We state the end results formally as follows:

THEOREM 5.1 *Let the $n\times n$ polynomial $P(z)$ be given in* (2.1), *and let $\lambda_1,\dots,\lambda_t$ be the distinct singular values for $P(z)$. Let $u_i(\lambda_k,\alpha)$, $v_i(\lambda_k,\alpha)$, $i=1,\dots,\sigma_k(\alpha)$, $\alpha=1,\dots,\tau_k$ be a complete normalized set of singular row and column sequences of $P(z)$ belonging to λ_k, $k=1,\dots,t$. If P_m is singular, then the principal part of $P(z)^{-1}$ at λ_k is*

$$\pi_k(z)=\sum_{\alpha=1}^{\tau_k}\sum_{j=1}^{\sigma_k(\alpha)}\frac{1}{(z-\lambda_k)^j}W_j(\lambda_k,\alpha),$$

where

$$W_j(\lambda_k,\alpha)=\sum_{h=1}^{\sigma_k(\alpha)+1-j} v_h(\lambda_k,\alpha)u_{\sigma_k(\alpha)+2-h-j}(\lambda_k,\alpha).$$

If P_m is nonsingular, then

$$P(z)^{-1}=\sum_{k=1}^{t}\pi_k(z).$$

REFERENCES

1 K. A. Abgarjan, Reduction of a square matrix to quasidiagonal form and its decomposition into components, *Izv. Akad. Nauk Armjan. SSR, Ser. Fiz.-Mat. Nauk* **18** (1965), 3–14.
2 T. Downs, On the inversion of a matrix of rational functions, *Linear Algebra Appl.* **4** (1971), 1–10.
3 E. Emre, Ö. Hüseyin, and K. Abdulla, On the inversion of rational matrices, *IEEE Trans. Circuits Syst.* **CAS-21** (1974), 8–10.
4 F. R. Gantmacher, *The Theory of Matrices*, Vol. 1, Chelsea, New York, 1960.
5 C. J. Hegedüs, Laurent expansion of the inverse of a function matrix, *Period. Math. Hung.* **6** (1975), 75–86.
6 R. L. Koust, The determination of the system transfer function matrix for flight control systems, *IEEE Trans.* **AC-13** (1968), 215.
7 C. E. Langenhop, The Laurent expansion for a nearly singular matrix, *Linear Algebra Appl.* **4** (1971), 329–340.
8 Carl E. Langenhop, On the invertibility of a nearly singular matrix, *Linear Algebra Appl.* **7** (1973), 361–365.

Received 1 October 1975; revised 21 May 1976