SIAM J. NUMER. ANAL.
Vol. 13, No. 6, December 1976

THE ALGEBRAIC THEORY OF MATRIX POLYNOMIALS*

J. E. DENNIS, JR.,† J. F. TRAUB‡ AND R. P. WEBER§

Abstract. A matrix S is a solvent of the matrix polynomial $M(X)=A_0X^m+\cdots+A_m$ if $M(S)=0$ where A_i, X, S are square matrices. In this paper we develop the algebraic theory of matrix polynomials and solvents. We define division and interpolation, investigate the properties of block Vandermonde matrices, and define and study the existence of a complete set of solvents. We study the relation between the matrix polynomial problem and the lambda-matrix problem, which is to find a scalar λ such that $A_0\lambda^m+A_1\lambda^{m-1}+\cdots+A_m$ is singular.

In a future paper we extend Traub's algorithm for calculating zeros of scalar polynomials to matrix polynomials and establish global convergence properties of this algorithm for a class of matrix polynomials.

1. Introduction. Let $A_0, A_1, \cdots, A_m, X$ be $n\times n$ complex matrices. We say

$$M(X)=A_0X^m+A_1X^{m-1}+\cdots+A_m \tag{1.1}$$

is a *matrix polynomial*. A matrix S is a *right solvent* of $M(X)$ if

$$M(S)=0.$$

The terminology *right* solvent is explained below. For simplicity we sometimes refer to right solvents as *solvents*. We say a matrix W is a *weak solvent* of $M(X)$ if $M(W)$ is singular. If A_0 is singular, one can shift coordinates and reverse the order of the coefficients to get a related problem with a nonsingular leading coefficient. We will ignore such problems and deal primarily with the case where $M(X)$ is *monic* $(A_0\equiv I)$.

We are interested in algorithms for the calculation of solvents. Since rather little is known about the mathematical properties of matrix polynomials and solvents, we develop that theory here. We extend division and interpolatory representation to matrix polynomials and study the properties of block Vandermonde matrices. In a future paper [3] we shall show that a generalization of Traub's algorithm [11] for calculating zeros of a scalar polynomial provides a globally convergent algorithm for calculating solvents for a class of matrix polynomials. We shall also report elsewhere on the use of solvents to solve systems of polynomial equations. Most of the results of this paper as well as additional material first appeared as a Carnegie-Mellon University-Cornell University Technical Report [2].

If the A_i are scalar matrices, $A_i=a_iI$, then (1.1) reduces to

$$M(X)=a_0X^m+a_1X^{m-1}+\cdots+a_m. \tag{1.2}$$

* Received by the editors December 24, 1973, and in final revised form November 12, 1975.

† Computer Science Department, Cornell University, Ithaca, New York 14853. The work of this author was supported in part by the National Science Foundation under Grant GJ27528.

‡ Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213. The work of this author was supported in part by the National Science Foundation under Grant GJ32111 and in part by the Office of Naval Research under Contract N0014-67-A-0314-0010, NR 044-422.

§ Bell Laboratories, Holmdel, New Jersey 07733. This work forms a portion of this author's Ph.D. thesis, Computer Science Department, Cornell University.

This problem has been thoroughly studied (Gantmacher, [4]) and we have such classical results as the Cayley–Hamilton theorem and the Lagrange–Sylvester interpolation theorem.

If X is a scalar matrix, $X = \lambda I$, then (1.1) reduces to

$$M(\lambda I) = M(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \cdots + A_m. \tag{1.3}$$

This is called a lambda-matrix and has also been thoroughly studied (Lancaster [7]). *Unfortunately, both* (1.2) *and* (1.3) *are sometimes called matrix polynomials but we shall reserve this name for* (1.1). A great deal is, of course, known about another special case of (1.1), the case of scalar polynomials ($n = 1$). A discussion of much of what is known about matrix polynomials may be found in MacDuffee [9]. The special case of calculating the square root of a matrix has been analyzed (Gantmacher [4]). Bell [1] studies the conditions under which an infinitude of solutions of (1.1) exist.

A problem closely related to that of finding solvents of a matrix polynomial is finding a scalar λ such that the lambda-matrix $M(\lambda)$ is singular. Such a scalar is called a *latent root* of $M(\lambda)$ and vectors $\mathbf{b}$ and $\mathbf{r}$ are right and left latent vectors, respectively, if for a latent root ρ, $M(\rho)\mathbf{b} = \mathbf{0}$ and $\mathbf{r}^T M(\rho) = \mathbf{0}^T$. See Lancaster [7], Gantmacher [4], MacDuffee [9], and Peters and Wilkinson [10] for discussions of latent roots.

The following relation between latent roots and solvents is well known (Lancaster [7]). A corollary of the generalized Bézout theorem states that if S is a solvent of $M(X)$, then

$$M(\lambda) = Q(\lambda)(I\lambda - S), \tag{1.4}$$

where $Q(\lambda)$ is a lambda-matrix of degree $m-1$. It is because of (1.4) that S is called a *right* solvent. The lambda-matrix $M(\lambda)$ has mn latent roots. From (1.4) the n eigenvalues of a solvent of $M(X)$ are all latent roots of $M(\lambda)$. The $n(m-1)$ latent roots of $Q(\lambda)$ are also latent roots of $M(\lambda)$. Thus if one is interested in the solution of a lambda-matrix problem, then a solvent will provide n latent roots and can be used for matrix deflation, which yields a new problem $Q(\lambda)$.

We summarize the results of this paper. In § 2 we define division for matrix polynomials and derive some important consequences. This definition includes scalar division and the generalized Bézout theorem as special cases. Section 3 contains a brief discussion of block companion matrices. In § 4 we give a sufficient condition for a matrix polynomial to have a complete set of solvents. In § 5 we introduce *fundamental matrix polynomials* and give a generalization of interpolation. In the final section we study the block Vandermonde matrix.

2. Division of matrix polynomials. We define division for matrix polynomials so that the class is closed under the operation. It reduces to scalar division if $n = 1$.

THEOREM 2.1. *Let* $M(X) = X^m + A_1X^{m-1} + \cdots + A_m$ *and* $W(X) = X^p + B_1X^{p-1} + \cdots + B_p$, *with* $m \geqq p$. *Then there exists a unique, monic matrix polynomial* $F(X)$ *of degree* $m-p$ *and a unique matrix polynomial* $L(X)$ *of degree not exceeding* $p-1$ *such that*

$$M(X) \equiv F(X)X^p + B_1F(X)X^{p-1} + \cdots + B_pF(X) + L(X). \tag{2.1}$$

Proof. Let $F(X) = X^{m-p} + F_1 X^{m-p-1} + \cdots + F_{m-p}$ and $L(X) = L_0 X^{p-1} + L_1 X^{p-2} + \cdots + L_{p-1}$. Equating coefficients of (2.1), $F_1, F_2, \cdots, F_{m-p}$ and $L_0, L_1, \cdots, L_{p-1}$ can be successively and uniquely determined from the m equations. Q.E.D.

Equation (2.1) is the *matrix polynomial division* of $M(X)$ on the left by $W(X)$ with quotient $F(X)$ and remainder $L(X)$.

DEFINITION 2.1. Associated with the matrix polynomial, $M(X) \equiv X^m + A_1 X^{m-1} + \cdots + A_m$, is the *commuted matrix polynomial*

$$\hat{M}(X) \equiv X^m + X^{m-1} A_1 + \cdots + A_m. \tag{2.2}$$

If $\hat{M}(R) = 0$, then R is a *left solvent* of $M(X)$.

An important association between the remainder, $L(X)$, and the dividend, $M(X)$, in (2.1) will now be given. It generalizes the fact that for scalar polynomials the dividend and remainder are equal when evaluated at the roots of the divisor.

COROLLARY 2.1. *If R is a left solvent of $W(X)$, then $\hat{L}(R) = \hat{M}(R)$.*

Proof. Let $Q(X) = M(X) - L(X)$. Then it is easily shown that

$$\hat{Q}(X) \equiv X^{m-p} \hat{W}(X) + X^{m-p-1} \hat{W}(X) F_1 + \cdots + \hat{W}(X) F_{m-p}. \tag{2.3}$$

The result then follows immediately since $\hat{Q}(R) = 0$ for all left solvents of $W(X)$. Q.E.D.

The case where $p = 1$ in Theorem 2.1 is of special importance in this paper. Here we have $W(X) = X - R$ where R is both a left and right solvent of $W(X)$. Then Theorem 2.1 shows that

$$M(X) \equiv F(X)X - RF(X) + L, \tag{2.4}$$

where L is a constant matrix. Now Corollary 2.1 shows that $L = \hat{M}(R)$, and thus

$$M(X) \equiv F(X)X - RF(X) + \hat{M}(R). \tag{2.5}$$

There is a corresponding theory for $\hat{M}(X)$. In this case, (2.1) is replaced by

$$\hat{M}(X) \equiv X^p \hat{H}(X) + X^{p-1} \hat{H}(X) B_1 + \cdots + \hat{H}(X) B_p + \hat{N}(X), \tag{2.6}$$

and Corollary 2.1 becomes the following.

COROLLARY 2.2. *If S is a right solvent of $W(X)$, then $N(S) = M(S)$.*

We again consider the case of $p = 1$. Let $W(X) = X - S$. Then (2.5) becomes

$$\hat{M}(X) \equiv X\hat{H}(X) - \hat{H}(X)S + M(S). \tag{2.7}$$

Restricting X to a scalar matrix λI, and noting that $M(\lambda) \equiv \hat{M}(\lambda)$, we get the generalized Bézout theorem (see Gantmacher [4, Chap. 4]) from (2.5) and (2.7):

$$M(\lambda) \equiv (I\lambda - R)F(\lambda) + \hat{M}(R) \equiv H(\lambda)(I\lambda - S) + M(S) \tag{2.8}$$

for any matrices R and S. If in addition R and S are left and right solvents, respectively, of $M(X)$, then

$$M(X) \equiv F(X)X - RF(X), \tag{2.9}$$

$$\hat{M}(X) \equiv X\hat{H}(X) - \hat{H}(X)S \tag{2.10}$$

and

$$M(\lambda) \equiv (I\lambda - R)F(\lambda) \equiv H(\lambda)(I\lambda - S). \tag{2.11}$$

Hence Corollaries 2.1 and 2.2 are generalizations of the generalized Bézout theorem.

Equation (2.11) *is the reason why R and S are called left and right solvents, respectively.*

The use of block matrices is fundamental in this work. It is useful to have a notation for the transpose of a matrix with square blocks.

DEFINITION 2.2 Let A be a matrix with block matrices (B_{ij}) of order n. The *block transpose* of dimension n of A, denoted $A^{B(n)}$, is the matrix with block matrices (B_{ji}).

The order of the block transpose will generally be dropped when it is clear. Note that, in general, $A^{B(n)} \neq A^T$, except that $n = 1$.

An important block matrix, which will be studied later in this paper, is the block Vandermonde matrix.

DEFINITION 2.3. Given $n \times n$ matrices $S_1, \cdots, S_m$, the *block Vandermonde* matrix is

$$V(S_1, \cdots, S_m) = \begin{pmatrix} I & I & \cdots & I \\ S_1 & S_2 & \cdots & S_m \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ S_1^{m-1} & S_2^{m-1} & \cdots & S_m^{m-1} \end{pmatrix}. \tag{2.12}$$

A scalar polynomial exactly divides another scalar polynomial, if all the roots of the divisor are also roots of the dividend. A generalization of the scalar polynomial result is given next. The notation is that of Theorem 2.1.

COROLLARY 2.3. *If $W(X)$ has p left solvents, $R_1, \cdots, R_p$, which are also left solvents of $M(X)$, and if $V^B(R_1, \cdots, R_p)$ is nonsingular, then the remainder $L(X) \equiv 0$.*

Proof. Corollary 2.1 shows that $\hat{L}(R_i) = 0$ for $i = 1, \cdots, p$. Since $V^B(R_1, \cdots, R_p)$ is nonsingular, and since

$$\begin{pmatrix} I & R_1 & \cdots & R_1^{p-1} \\ I & R_2 & \cdots & R_2^{p-1} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ I & R_p & \cdots & R_p^{p-1} \end{pmatrix} \begin{pmatrix} L_{p-1} \\ L_{p-2} \\ \cdot \\ \cdot \\ \cdot \\ L_0 \end{pmatrix} = \begin{pmatrix} \hat{L}(R_1) \\ \hat{L}(R_2) \\ \cdot \\ \cdot \\ \cdot \\ \hat{L}(R_p) \end{pmatrix} = 0$$

it follows that $L(X) \equiv 0$. Thus

$$M(X) \equiv F(X)X^p + B_1F(X)X^{p-1} + \cdots + B_pF(X). \tag{2.13}$$

Q.E.D.

From (2.11) it follows that the eigenvalues of any solvent (left or right) of $M(X)$ are latent roots of $M(\lambda)$. These equations allow us to think of right (left) solvents of $M(X)$ as right (left) factors of $M(\lambda)$.

In the scalar polynomial case, due to commutivity, right and left factors are equivalent. Relations between left and right solvents can now be given.

COROLLARY 2.4. *If S and R are right and left solvents of $M(X)$, respectively, and S and R have no common eigenvalues, then $F(S)=0$, where $F(X)$ is the quotient of division of $M(X)$ on the left by $X-R$* (see equation (2.9)).

Proof. Equation (2.9) shows that

$$F(S)S-RF(S)=0. \tag{2.14}$$

Since S and R have no common eigenvalues, $F(S)=0$ uniquely. This follows, since the solution of $AX=XB$ has the unique solution $X=0$, if and only if A and B have no common eigenvalues. See Gantmacher [4, Chap. 8]. Q.E.D.

Given a left solvent R of $M(X)$, Theorem 2.1 shows that $F(X)$ exists uniquely. If S is a right solvent of $M(X)$ and if $F(S)$ is nonsingular (S is not a weak solvent of $F(X)$), then (2.14) shows that

$$R=F(S)SF(S)^{-1}. \tag{2.15}$$

This gives an association between left and right solvents.

3. Block companion matrix. A useful tool in the study of scalar polynomials is the companion matrix which permits us to bring matrix theory to bear on the analysis of polynomial zeros. We study properties of block companion matrices. Definition 3.1, Theorem 3.1 and Corollary 3.1 can be found in Lancaster [7].

DEFINITION 3.1. Given a matrix polynomial

$$M(X)\equiv X^m+A_1X^{m-1}+\cdots+A_m,$$

the *block companion matrix* associated with it is

$$C=\begin{pmatrix} 0 & \cdots & 0 & -A_m \\ I & & & -A_{m-1} \\ & \ddots & & \vdots \\ & & I & -A_1 \end{pmatrix}. \tag{3.1}$$

It is well known that the eigenvalues of the block companion matrix are latent roots of the associated lambda-matrix. Simple algebraic manipulation yields

THEOREM 3.1. $\det(C-\lambda I)\equiv(-1)^{mn}\det(I\lambda^m+A_1\lambda^{m-1}+\cdots+A_m)$.

Since C is an mn by mn matrix, we immediately obtain the following well known result.

COROLLARY 3.1. *$M(\lambda)$ has exactly mn finite latent roots.*

Note that if $M(\lambda)$ were not monic and had a singular leading coefficient, then the lambda-matrix can be viewed as having latent roots at infinity.

The form of the block companion matrix could have been chosen differently. Theorem 3.1 also holds for the block transpose of the companion matrix:

$$C^B = \begin{pmatrix} 0 & I & & \\ \cdot & & \cdot & \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ 0 & & & I \\ -A_m & -A_{m-1} & \cdots & -A_1 \end{pmatrix}. \tag{3.2}$$

It will be useful to know the eigenvectors of the block companion matrix and its block transpose. The results are a direct generalization of the scalar case and are easily verified. See for example Jenkins and Traub [6] or Lancaster and Webber [8].

THEOREM 3.2. *If ρ_i is a latent root of $M(\lambda)$ and $\mathbf{b}_i$ and $\mathbf{r}_i$ are right and left latent vectors, then ρ_i is an eigenvalue of C and of C^B and*

(3.3)

(i) $\begin{pmatrix} \mathbf{b}_i \\ \rho_i \mathbf{b}_i \\ \cdot \\ \cdot \\ \cdot \\ \rho_i^{m-1}\mathbf{b}_i \end{pmatrix}$ *is the right eigenvector of C^B,*

(ii) $\begin{pmatrix} \mathbf{r}_i \\ \rho_i \mathbf{r}_i \\ \cdot \\ \cdot \\ \cdot \\ \rho_i^{m-1}\mathbf{r}_i \end{pmatrix}$ *is the left eigenvector of C, and*

(iii) $\begin{pmatrix} \mathbf{b}_i^{(m-1)} \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{b}_i^{(1)} \\ \mathbf{b}_i \end{pmatrix}$ *is the right eigenvector of C,*

where

$$\frac{M(\lambda)\mathbf{b}_i}{\lambda - \rho_i} \equiv \mathbf{b}_i \lambda^{m-1} + \mathbf{b}_i^{(1)} \lambda^{m-2} + \cdots + \mathbf{b}_i^{(m-1)}.$$

4. Structure and existence of solvents. We introduce the concept of a "complete set" of solvents. This is analogous to the fact that a scalar polynomial of degree n has n zeros. We establish a sufficient condition for a matrix polynomial to have a complete set of solvents.

The fundamental theorem of algebra does not hold for matrix polynomials. This is known from the extensive studies of the square-root problem: $X^2 = A$. See Gantmacher [4, p. 231].

A sufficient condition for the existence of a solvent (Lancaster [7, p. 49]) is given by

LEMMA 4.1. *If $M(\lambda)$ has n linearly independent right latent vectors, $\mathbf{b}_1, \cdots, \mathbf{b}_n$, corresponding to latent roots $\rho_1, \cdots, \rho_n$, then $Q\Lambda Q^{-1}$ is a right solvent, where $Q = [\mathbf{b}_1, \cdots, \mathbf{b}_n]$ and $\Lambda = \operatorname{diag}(\rho_1, \cdots, \rho_n)$.*

It follows from the above theorem that if a solvent is diagonalizable, then it must be the form $Q\Lambda Q^{-1}$ where the columns of Q are right latent vectors of $M(\lambda)$.

COROLLARY 4.1. *If $M(\lambda)$ has mn distinct latent roots, and the set of right latent vectors satisfy the Haar condition (that every set of n of them are linearly independent), then there are exactly $\binom{mn}{n}$ different right solvents.*

Consider now the special case of a matrix polynomial whose associated lambda-matrix has distinct roots. We call a set of m solvents a *complete set* if the mn eigenvalues of this set exactly match the distinct latent roots. Note that we have defined the concept of complete set only for the case of distinct latent roots. In Theorem 4.1 we shall show that in this case a complete set of solvents exists. We consider the following example to illustrate Lemma 4.1, the definition of a complete set of solvents and Theorem 4.1.

Consider the quadratic

$$M(X) = X^2 + \begin{pmatrix} -1 & -6 \\ 2 & -9 \end{pmatrix} X + \begin{pmatrix} 0 & 12 \\ -2 & 14 \end{pmatrix}.$$

The corresponding lambda-matrix has latent roots 1, 2, 3, 4 with corresponding latent vectors $(1, 0)^T$, $(0, 1)^T$, $(1, 1)^T$, $(1, 1)^T$. The problem has a complete set of solvents $S_1 = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ and $S_2 = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$. Other solvents have eigenvalues 1, 2; 1, 4 and 2, 3. The only pair which cannot be the eigenvalues of a solvent is 3, 4.

Before proceeding with Theorem 4.1 we prove two preliminary results.

LEMMA 4.2. *If a matrix A is nonsingular, then there exists a permutation of the columns of A to $\tilde{A}$ such that $\tilde{A} = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{pmatrix}$, where $\tilde{A}_{11}$ is square and of arbitrary order less than n and $\tilde{A}_{11}$ and $\tilde{A}_{22}$ are nonsingular.*

Proof. Let $1 \leqq k < n$. Expand $\det(A)$ in terms of square matrices formed from the first k rows using the Laplace expansion. Since $\det(A) \neq 0$, one of the products in the expansion is nonzero and the result follows. Q.E.D.

Once the columns of A are permuted to get $\tilde{A}_{11}$ and $\tilde{A}_{22}$ nonsingular, the process can be continued to similarly divide $\tilde{A}_{22}$ into nonsingular blocks without destroying the nonsingularity of $\tilde{A}_{11}$. We thus arrive at

LEMMA 4.3. *If A, a matrix of order mn, is nonsingular, then there exists a permutation of the columns of A to $\tilde{A} = (B_{ij})$, with B_{ij} a matrix of order n, such that B_{ii} is nonsingular for $i = 1, \cdots, m$.*

The existence theorem is given by

THEOREM 4.1. *If the latent roots of $M(\lambda)$ are distinct, then $M(X)$ has a complete set of solvents.*

Proof. If the latent roots of $M(\lambda)$ are distinct, then the eigenvalues of the block companion matrix are distinct, and hence the eigenvectors of the block companion matrix are linearly independent. From Theorem 3.2 the set of vectors

$$\begin{pmatrix} \mathbf{b}_i \\ \rho_i \mathbf{b}_i \\ \cdot \\ \cdot \\ \cdot \\ \rho_i^{m-1} \mathbf{b}_i \end{pmatrix},$$

for $i = 1, \cdots, mn$, are eigenvectors of C^B. The matrix whose columns are these mn vectors is nonsingular. Lemma 4.3 shows that there are m disjoint sets of n linearly independent vectors $\mathbf{b}_i$. Using the structure $Q\Lambda Q^{-1}$ of Lemma 4.1, the complete set of solvents can be formed. Q.E.D.

COROLLARY 4.2. *If $M(\lambda)$ has distinct latent roots, then it can be factored into the product of linear lambda-matrices.*

Proof. Since $M(\lambda)$ has distinct latent roots, there exists a right solvent S and $M(\lambda) = Q(\lambda)(I\lambda - S)$. $Q(\lambda)$ has the remaining latent roots of $M(\lambda)$ as its latent roots. It follows then, that the latent roots of $Q(\lambda)$ are distinct. Thus the process can be continued until the last quotient is linear. Q.E.D.

The process described in the above proof considers solvents of the sequence of lambda-matrices formed by the division $M(\lambda) = Q(\lambda)(I\lambda - S)$.

DEFINITION 4.1. A sequence of matrices $C_1, \cdots, C_m$ form a *chain of solvents of* $M(X)$ if C_i is a right solvent of $Q_i(X)$, where $Q_m(X) \equiv M(x)$ and

$$Q_i(\lambda) \equiv Q_{i-1}(\lambda)(I\lambda - C_i), \qquad i = m, \cdots, 1. \tag{4.1}$$

It should be noted that, in general, only C_m is a right solvent of $M(X)$. Furthermore, C_1 is a left solvent of $M(X)$. An equivalent definition of a chain of solvents could be defined with C_i, a left solvent of $T_i(X)$, and

$$T_i(\lambda) \equiv (I\lambda - C_{m-i+1})T_{i-1}(\lambda), \qquad i = m, \cdots, 1. \tag{4.2}$$

COROLLARY 4.3. *If the latent roots of $M(\lambda)$ are distinct, then $M(X)$ has a chain of solvents.*

If $C_1, \cdots, C_m$ form a chain of solvents of $M(X)$, then

$$M(\lambda) \equiv I\lambda^m + A_1\lambda^{m-1} + \cdots + A_m \equiv (I\lambda - C_1)(I\lambda - C_2)\cdots(I\lambda - C_m). \tag{4.3}$$

This leads to a generalization of the classical result for scalar polynomials which relates coefficients to elementary symmetric functions. By equating coefficients of (4.3), one gets the following

THEOREM 4.2. *If $C_1, \cdots, C_m$ form a chain of solvents for $M(X) = X^m + A_1X^{m-1} + \cdots + A_m$, then*

$$(4.4)\qquad \begin{aligned} A_1 &= -(C_1 + C_2 + \cdots + C_m),\\ A_2 &= (C_1C_2 + C_1C_3 + \cdots + C_{m-1}C_m),\\ &\vdots\\ A_m &= (-1)^m C_1C_2\cdots C_m. \end{aligned}$$

5. Interpolation and representation. Given scalars $s_1, \cdots, s_m$, the *fundamental polynomials* $m_i(x)$ of interpolation theory are defined so that $m_i(s_j) = \delta_{ij}$. We generalize these relations for our matrix problem.

DEFINITION 5.1. Given a set of matrices, $S_1, \cdots, S_m$, the *fundamental matrix polynomials* are a set of $m-1$ degree matrix polynomials, $M_1(X), \cdots, M_m(X)$, such that $M_i(S_j) = \delta_{ij}I$.

Sufficient conditions on the set of matrices, $S_1, \cdots, S_m$, for a set of fundamental matrix polynomials to exist uniquely will be given in Theorem 5.2. First, however, we need the following easily proven result.

THEOREM 5.1. *Given m pairs of matrices, (X_i, Y_i), $i = 1, \cdots, m$, then there exists unique matrix polynomials*

$$P_1(X) = A_1X^{m-1} + A_2X^{m-2} + \cdots + A_m$$

and

$$P_2(X) = X^m + B_1X^{m-1} + \cdots + B_m$$

such that

$$P_1(X_i) = P_2(X_i) = Y_i$$

for $i = 1, \cdots, m$ if and only if $V(X_1, \cdots, X_m)$ is nonsingular.

Let $M(X)$ have a complete set of solvents, $S_1, \cdots, S_m$, such that $V(S_1, \cdots, S_m)$ is nonsingular. According to Theorem 5.1, there exists a unique matrix polynomial

$$(5.1)\qquad M_i(X) \equiv A_1^{(i)}X^{m-1} + \cdots + A_m^{(i)}$$

such that

$$(5.2)\qquad M_i(S_j) = \delta_{ij}I.$$

Note that $M_i(X)$ has the same solvents as $M(X)$, except S_i has been deflated out. The $M_i(X)$ are the fundamental matrix polynomials.

Denote by $V(S_1, \cdots, S_{i-1}, S_{i+1}, \cdots, S_m)$ the block Vandermonde at the $m-1$ solvents, $S_1, \cdots, S_m$, with S_i deleted.

THEOREM 5.2. *If matrices $S_1, \cdots, S_m$ are such that $V(S_1, \cdots, S_m)$ is nonsingular, then there exist unique matrix polynomials $M_i(X) \equiv A_1^{(i)}X^{m-1} + \cdots + A_m^{(i)}$ for $i = 1, \cdots, m$ such that $M_1(X), \cdots, M_m(X)$ are* *fundamental matrix polynomials. If, furthermore,* $V(S_1, \cdots, S_{k-1}, S_{k+1}, \cdots, S_m)$ *is nonsingular, then* $A_1^{(k)}$ *is nonsingular.*

Proof. $V(S_1, \cdots, S_m)$ nonsingular implies that there exists a unique set of fundamental matrix polynomials, $M_1(X), \cdots, M_m(X)$. $V(S_1, \cdots, S_{k-1}, S_{k+1}, \cdots, S_m)$ nonsingular and Theorem 5.1 imply that there exists a unique monic matrix polynomial $N_k(X) \equiv X^{m-1} + N_1^{(k)}X^{m-2} + \cdots + N_m^{(k)}$, such that $N_k(S_j) = 0$ for $j \neq k$. Consider $Q_k(X) \equiv N_k(S_k)M_k(X)$. $Q_k(S_j) = N_k(S_j)$ for $j = 1, \cdots, m$. Since $V(S_1, \cdots, S_m)$ is nonsingular and both $Q_k(X)$ and $N_k(X)$ are of degree $m-1$, it follows that $Q_k(X) \equiv N_k(X)$. Thus $N_k(X) \equiv N_k(S_k)M_k(X)$. Equating leading coefficients, we get $I = N_k(S_k)A_1^{(k)}$, and thus $A_1^{(k)}$ is nonsingular. Q.E.D.

The fundamental matrix polynomials, $M_1(X), \cdots, M_m(X)$, can be used in a generalized Lagrange interpolation formula. Paralleling the scalar case, we get the following representation theorems.

THEOREM 5.3. *If matrices* $S_1, \cdots, S_m$ *are such that* $V(S_1, \cdots, S_m)$ *is nonsingular, and* $M_1(X), \cdots, M_m(X)$ *are a set of fundamental matrix polynomials, then, for an arbitrary*

$$G(X) \equiv B_1X^{m-1} + \cdots + B_m, \tag{5.3}$$

it follows that

$$G(X) = \sum_{i=1}^{m} G(S_i)M_i(X). \tag{5.4}$$

Proof. Let $Q(X) = \sum_{i=1}^{m} G(S_i)M_i(X)$. Then $Q(S_i) = G(S_i)$ for $i = 1, \cdots, m$. Since the block Vandermonde is nonsingular, it follows that $Q(X)$ is unique and hence $G(X) \equiv Q(X)$. Q.E.D.

A lambda-matrix was defined as a matrix polynomial whose variable was restricted to the scalar matrix λI. Thus the previous theorem holds for lambda-matrices as well.

COROLLARY 5.1. *Under the same assumptions as in Theorem* 5.3, *for an arbitrary lambda-matrix*

$$G(\lambda) \equiv B_1\lambda^{m-1} + \cdots + B_m, \tag{5.5}$$

it follows that

$$G(\lambda) = \sum_{i=1}^{m} G(S_i)M_i(\lambda). \tag{5.6}$$

Fundamental matrix polynomials were defined such that $M_i(S_j) = \delta_{ij}I$. A result similar to (2.9) can be derived based on the fundamental matrix polynomials. It was previously (§ 2) developed using matrix polynomial division.

THEOREM 5.4. *If* $M(X)$ *has a set of right solvents,* $S_1, \cdots, S_m$, *such that* $V(S_1, \cdots, S_m)$ *and* $V(S_1, \cdots, S_{i-1}, S_{i+1}, \cdots, S_m)$ *for each* $i = 1, \cdots, m$ *are nonsingular and* $M_1(X), \cdots, M_m(X)$ *are the set of fundamental matrix polynomials, then*

$$M_i(X)X - S_iM_i(X) \equiv A_1^{(i)}M(X) \quad \text{for } i = 1, \cdots, m, \tag{5.7}$$

where $A_1^{(i)}$ *is the leading matrix coefficient of* $M_i(X)$.

Proof. Let $Q_i(X) \equiv M_i(X)X - S_iM_i(X)$. Note that $Q_i(S_j)=0$ for all j. $M(X)$ is the unique monic matrix polynomial with right solvents $S_1, \cdots, S_m$ since $V(S_1, \cdots, S_m)$ is nonsingular. The leading matrix coefficient of $Q_i(X)$ is $A_1^{(i)}$ which is nonsingular, since $V(S_1, \cdots, S_{i-1}, S_{i+1}, \cdots, S_m)$ is nonsingular. Thus $M(X) \equiv A_1^{(i)^{-1}}Q_i(X)$. Q.E.D.

A previous result (equation (2.5)) stated that if R_i is a left solvent of $M(X)$, then there exists a unique, monic polynomial $F_i(X)$ of degree $m-1$ such that

$$M(X) \equiv F_i(X)X - R_iF_i(X). \tag{5.8}$$

Comparing (5.7) and (5.8), we obtain the following result.

COROLLARY 5.2. *Under the conditions of Theorem* 5.4, $F_i(X) \equiv [A_1^{(i)}]^{-1}M_i(X)$ *and*

$$R_i = [A_1^{(i)}]^{-1}S_iA_1^{(i)} \tag{5.9}$$

is a left solvent of $M(X)$.

If $M(X)$ has a set of right solvents, $S_1, \cdots, S_m$, such that $V(S_1, \cdots, S_m)$ and $V(S_1, \cdots, S_{i-1}, S_{i+1}, \cdots, S_m)$ for $i = 1, \cdots, m$ are all nonsingular, then by (5.9), there exists a set of left solvents of $M(X)$, $R_1, \cdots, R_m$, such that R_i is similar to S_i for all i.

COROLLARY 5.3. *Under the conditions of Theorem* 5.4, *if* R_i *is defined as in* (5.9), *then*

$$\bar{M}_i(\lambda) \equiv [A_1^{(i)}]^{-1}M_i(\lambda) = (I\lambda - R_i)^{-1}M(\lambda). \tag{5.10}$$

Proof. The result follows from (5.8) and Corollary 5.2. Q.E.D.

6. Block Vandermonde. The block Vandermonde matrix is of fundamental importance to the theory of matrix polynomials. This section considers some of its properties.

It is well known that in the scalar case ($n = 1$),

$$\det V(s_1, \cdots, s_m) = \prod_{i>j} (s_i - s_j), \tag{6.1}$$

and thus the Vandermonde is nonsingular if the set of s_i's are distinct. One might expect that if the eigenvalues of X_1 and X_2 are disjoint and distinct, then $V(X_1, X_2)$ is nonsingular. That this is not the case is shown by the following example. The determinant of the block Vandermonde at two points is

$$\det V(X_1, X_2) = \det \begin{pmatrix} I & I \\ X_1 & X_2 \end{pmatrix} = \det (X_2 - X_1). \tag{6.2}$$

Even if X_1 and X_2 have no eigenvalues in common, $X_2 - X_1$ may still be singular. The example $X_1 = \begin{pmatrix} 2 & 0 \\ -2 & 1 \end{pmatrix}$ and $X_2 = \begin{pmatrix} 4 & 2 \\ 0 & 3 \end{pmatrix}$ yields $X_2 - X_1$ singular. Now if we include $X_3 = \begin{pmatrix} 6 & 2 \\ 0 & 5 \end{pmatrix}$, we find $V(X_1, X_2, X_3)$ to be nonsingular. Thus it is possible for $V(X_1, X_2, X_3)$ to be nonsingular and $V(X_1, X_2)$ to be singular.

It will be shown that the X_1 and X_2 in this example cannot be the set of solvents of a monic quadratic matrix polynomial. First we state a well known property of an invariant subspace.

LEMMA 6.1. *Let matrix A have distinct eigenvalues and N be a subspace of C^n of dimension d. Suppose further that if $v \in N$, then $A\mathbf{v} \in N$. Under these conditions, d of the eigenvectors of A are in N.*

Our main result of this section can now be stated:

THEOREM 6.1. *If $M(\lambda)$ has distinct latent roots, then there exists a complete set of right solvents of $M(X)$, $S_1, \cdots, S_m$, and for any such set of solvents, $V(S_1, \cdots, S_m)$ is nonsingular.*

Proof. The existence was proved in Theorem 4.1. The hypothesis that $S_1, \cdots, S_m$, are right solvents of $M(X) = X^m + A_1X^{m-1} + \cdots + A_m$, is equivalent to

$$(6.3) \qquad (A_m, \cdots, A_1)\begin{pmatrix} I & \cdots & I \\ S_1 & \cdots & S_m \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ S_1^{m-1} & \cdots & S_m^{m-1} \end{pmatrix} = (-S_1^m, \cdots, -S_m^m).$$

Assume $\det V(S_1, \cdots, S_m) = 0$ and let N be the nullspace of $V(S_1, \cdots, S_m)$. Letting $\mathbf{v} \in N$ and using (6.3), we get

$$\mathbf{0} = (A_m, \cdots, A_1)\, V(S_1, \cdots, S_m)\mathbf{v} = (-S_1^m, \cdots, -S_m^m)\mathbf{v}.$$

Hence for all $\mathbf{v} \in N$, we have

$$(6.4) \qquad \mathbf{0} = \begin{pmatrix} S_1 & \cdots & S_m \\ S_1^2 & \cdots & S_m^2 \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ S_1^m & \cdots & S_m^m \end{pmatrix}\mathbf{v} = V(S_1, \cdots, S_m)\operatorname{diag}(S_1, \cdots, S_m)\mathbf{v}.$$

Letting $D = \operatorname{diag}(S_1, \cdots, S_m)$, (6.4) shows that for all $\mathbf{v} \in N$, $D\mathbf{v} \in N$. Since D has distinct eigenvalues, Lemma 6.1 applies, and there are as many eigenvectors of D in N as the dimension of N. Since the eigenvalues of the solvents were assumed to be distinct, it follows that the eigenvectors of D are of the form $(\mathbf{0}^T, \cdots, \mathbf{0}^T, \mathbf{w}^T, \mathbf{0}^T, \cdots, \mathbf{0}^T)$, where $\mathbf{w}$ is an eigenvector of one of the S_i's. Let $\mathbf{u} = (\mathbf{0}^T, \cdots, \mathbf{0}^T, \mathbf{w}^T, \mathbf{0}^T, \cdots, \mathbf{0}^T)$ be an arbitrary eigenvector of D in N. Thus $V(S_1, \cdots, S_m)\mathbf{u} = \mathbf{0}$. But then, $I\mathbf{w} = \mathbf{0}$, which is a contradiction. Thus $\det V(S_1, \cdots, S_m) \neq 0$. Q.E.D.

The example considered before this theorem, was a case where matrices X_1 and X_2 had distinct and disjoint eigenvalues and $\det V(X_1, X_2) = 0$. Thus by the theorem, they could not be a set of right solvents for a monic, quadratic matrix polynomial. In contrast with the theory of scalar polynomials, we have the following result.

COROLLARY 6.1. *There exist sets containing m matrices which are not a set of right solvents for any monic matrix polynomial of degree m.*

We now tie Theorem 6.1 together with the results of § 5 to obtain the following result.

THEOREM 6.2. *If $M(\lambda)$ has distinct latent roots, then*

1. *there exists a complete set of solvents, $S_1, \cdots, S_m$,*
2. *$V(S_1, \cdots, S_m)$ is nonsingular,*
3. *there exists a set of fundamental matrix polynomials, $M_i(X)$, such that $M_i(S_j) = \delta_{ij}I$, and*
4. *for an arbitrary $G(X)$,*

$$G(X) = \sum_{i=1}^{m} G(S_i)M_i(X).$$

We now prove a generalization of (6.1), that the Vandermonde of scalars is the product of the differences of the scalars. Let $M^{(d)}_{S_1\cdots S_k}(X)$ be a monic matrix polynomial of degree $d \geqq k$ with right solvents $S_1, \cdots, S_k$. The superscript d will be omitted if $d = K$.

THEOREM 6.3. *If $V(S_1, \cdots, S_k)$ is nonsingular for $k = 2, \cdots, r-1$, then*

$$\det V(S_1, \cdots, S_r) = \det V(S_1, \cdots, S_{r-1}) \det M_{S_1\cdots S_{r-1}}(S_r). \tag{6.5}$$

Proof. The nonsingularity of $V(S_1, \cdots, S_{r-1})$ and Theorem 5.1 guarantee that $M_{S_1\cdots S_{r-1}}(X)$ exists uniquely. The determinant of $V(S_1, \cdots, S_r)$ will be evaluated by block Gaussian elimination using the fact that for an arbitrary matrix E of the proper dimension,

$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det\begin{pmatrix} A+EC & B+ED \\ C & D \end{pmatrix}. \tag{6.6}$$

$$\det V(S_1, \cdots, S_r) = \det\begin{pmatrix} I & \cdots & I \\ S_1 & \cdots & S_r \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ S_1^{r-1} & \cdots & S_r^{r-1} \end{pmatrix}$$

$$= \det\begin{pmatrix} I & I & \cdots & I \\ & S_2 - S_1 & \cdots & S_r - S_1 \\ & \cdot & & \cdot \\ & \cdot & & \cdot \\ & \cdot & & \cdot \\ & S_2^{r-1} - S_1^{r-1} & \cdots & S_r^{r-1} - S_1^{r-1} \end{pmatrix} \tag{6.7}$$

$$= \det \begin{pmatrix} I & I & I & \cdots & I \\ & S_2 - S_1 & S_3 - S_1 & \cdots & S_r - S_1 \\ & & M^{(2)}_{S_1S_2}(S_3) & \cdots & M^{(2)}_{S_1S_2}(S_r) \\ & & \vdots & & \vdots \\ & & M^{(r-1)}_{S_1S_2}(S_3) & \cdots & M^{(r-1)}_{S_1S_2}(S_r) \end{pmatrix}$$

where $M^{(d)}_{S_1S_2}(X) = (X^d - S_1^d) - (S_2^d - S_1^d)(S_2 - S_1)^{-1}(X - S_1)$. $(S_2 - S_1)$ is nonsingular, since $\det(S_2 - S_1) = \det V(S_1, S_2) \neq 0$. It will be shown that after k steps of the block Gaussian elimination, the general term for the i, j block, $i, j > k$, is $M^{(i-1)}_{S_1\cdots S_k}(S_j)$. Assume it is true after $k-1$ steps. Then after k steps, the i, j element is

$$M^{(i-1)}_{S_1\cdots S_{k-1}}(S_j) - M^{(i-1)}_{S_1\cdots S_{k-1}}(S_k) M^{(k-1)}_{S_1\cdots S_{k-1}}(S_k)^{-1} M^{(k-1)}_{S_1\cdots S_{k-1}}(S_j).$$

This is merely $M^{(i-1)}_{S_1\cdots S_k}(X)$ evaluated at $X = S_j$. Using the fact that the determinant of a block triangular matrix is the product of the determinants of the diagonal matrices (see Householder [5]), the result follows. Q.E.D.

COROLLARY 6.2. *If $V(S_1, \cdots, S_{k-1})$ is nonsingular and S_k is not a weak solvent of $M_{S_1\cdots S_{k-1}}(X)$, then $V(S_1, \cdots, S_k)$ is nonsingular.*

It is useful to be able to construct matrix polynomials with a given set of right solvents.

COROLLARY 6.3. *Given matrices $S_1, \cdots, S_m$ such that $V(S_1, \cdots, S_k)$ is nonsingular for $k = 2, \cdots, m$, the iteration $N_0(X) = I$,*

$$N_i(X) = N_{i-1}(X)X - N_{i-1}(S_i)S_iN_{i-1}^{-1}(S_i)N_{i-1}(X) \tag{6.8}$$

is defined and yields an m degree monic matrix polynomial $N_m(X)$, such that $N_m(S_i) = 0$ for $i = 1, \cdots, m$.

Proof. $N_1(X) \equiv X - S_1 = M_{S_1}(X)$. Assume $N_k(X) \equiv M_{S_1\cdots S_k}(X)$. Then from (6.8), $N_{k+1}(S_i) = 0$ for $i = 1, \cdots, k+1$ and hence $N_{k+1}(X) \equiv M_{S_1\cdots S_{k+1}}(X)$. The sequence of block Vandermondes being nonsingular guarantees the nonsingularity of $N_{i-1}(S_i)$. Q.E.D.

REFERENCES

[1] J. BELL, *Families of solutions of the unilateral matrix equation*, Proc. Amer. Math. Soc., 1 (1950) pp. 151–159.
[2] J. E. DENNIS, J. F. TRAUB AND R. P. WEBER, *On the matrix polynomial, lambda-matrix and block eigenvalue problems*, Computer Science Department tech. rep., Cornell Univ., Ithaca, N.Y., and Carnegie-Mellon Univ., Pittsburgh, Pa., 1971.
[3] J. E. DENNIS, J. F. TRAUB AND R. P. WEBER, *Algorithms for solvents of matrix polynomials*, submitted.
[4] F. R. GANTMACHER, *The Theory of Matrices. I, II*, Chelsea, New York, 1960.
[5] A. HOUSEHOLDER, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964.
[6] M. A. JENKINS AND J. F. TRAUB, *A three-stage variable-shift iteration for polynomial zeros and its relation to generalized Rayleigh iteration*, Numer. Math., 14 (1970), pp. 252–263.
[7] P. LANCASTER, *Lambda-Matrices and Vibrating Systems*, Pergamon Press, New York, 1966.

[8] P. LANCASTER AND P. WEBBER, *Jordan chains for lambda matrices*, Linear Algebra and Appl., 1 (1968), pp. 563–569.
[9] C. MACDUFFEE, *The Theory of Matrices*, Chelsea, New York, 1946.
[10] G. PETERS AND J. WILKINSON, $AX = \lambda BX$ *and the generalized eigenproblem*, this Journal, 7 (1970), pp. 479–492.
[11] J. F. TRAUB, *A class of globally convergent iteration functions for the solution of polynomial equations*, Math. Comp., 20 (1966), pp. 113–138.