SIAM J. MATRIX ANAL. APPL.
Vol. 13, No. 4, pp. 1123–1130, October 1992

008

EXPLICIT SOLUTIONS OF THE MATRIX EQUATION $\sum A^iXD_i = C$*

HARALD K. WIMMER†

Abstract. A module theoretic approach is developed to study the linear matrix equation $\sum A^iXD_i = C$. If the solution X is unique it can be expressed in the form $X = \sum A^kCG_k$. The matrices G_k can be determined from an auxiliary equation which involves the companion matrix of the characteristic polynomial of A. A connection is made with realizations of the inverse of the associated polynomial matrix $D(x) = \sum D_ix^i$. The more general equation $\sum A_iXD_i \doteq C$ is discussed under the assumption that the matrices A_i are pairwise commutative.

Key words. linear matrix equations, Sylvester equations, realizations, simultaneous triangular form

AMS(MOS) subject classification. 15A24

1. Introduction. Linear matrix equations have been studied since the time of Sylvester [10]. They are important in numerous applications (see, e.g., [5] with the regulator problem as a more recent example). An algebraic approach that originated with Kalman [7], [8] was elaborated by Djaferis and Mitter [1] to yield explicit solutions of equations of the form $\sum f_{ik}A^iXB^k = C$ (see also [12] and [13]). In this note, we adapt the module theoretic point of view of those papers and consider the more general equation

$$\sum_{i=0}^{p-1} A^iXD_i = C. \tag{1.1}$$

We first review some known facts concerning the equation

$$\sum_{i=0}^{p-1}\sum_{k=0}^{q-1} f_{ik}A^iXB^k = C, \tag{1.2}$$

which should direct us towards corresponding results on (1.1). All matrices in (1.1) and (1.2) are assumed to be complex, in particular, $A \in \mathbb{C}^{p\times p}$, B, $D_i \in \mathbb{C}^{q\times q}$, $C \in \mathbb{C}^{p\times q}$. Furthermore, the polynomial

$$f(x, y) = \sum\sum f_{ik}x^iy^k$$

is in $\mathbb{C}[x, y]$. For the results of this section, we refer to [1], [12], and [15].

Let $\lambda_1, \cdots, \lambda_p$ be the eigenvalue of A and let

$$a(x) = a_0 + \cdots + a_{p-1}x^{p-1} + x^p$$

be the characteristic polynomial of A. Define a companion matrix of $a(x)$ by

$$F_a = \begin{pmatrix} 0 & & & -a_0 \\ 1 & & & \\ & \ddots & & \vdots \\ & & 1 & -a_{p-1} \end{pmatrix}.$$

Let B have eigenvalues $\mu_1, \cdots, \mu_q$ and characteristic polynomial $b(x)$. The ideal in $\mathbb{C}[x, y]$ generated by $a(x)$ and $b(y)$ will be denoted by Ψ, i.e.,

$$\Psi = (a(x), b(y)).$$

---

* Received by the editors October 31, 1988; accepted for publication (in revised form) December 13, 1990.

† Mathematisches Institut, Universität Würzburg, D-8700 Würzburg, Germany.

Equation (1.2) is called *universally solvable* if it has a solution for every C (which is necessarily unique).

For $h = \sum h_{ik}x^iy^k \in \mathbb{C}[x, y]$ and $M \in \mathbb{C}^{p\times q}$, let a product $h * M$ be defined by

$$h * M = \sum_{i,k} h_{ik}A^iMB^k.$$

The Cayley–Hamilton theorem implies that $d * M = 0$ for all $d \in \Psi$. Hence it is meaningful to define

$$(H + \Psi) * M = h * M$$

such that $\mathbb{C}^{p\times q}$ becomes a module over the ring $\mathbb{C}[x, y]/\Psi$. Then (1.2) can be written as $(f + \Psi) * X = C$.

THEOREM 1.1. *The following statements are equivalent*:

(i) *Equation* (1.2) *is universally solvable.*

(ii) *The following holds*:

$$f(\lambda_i, \mu_j) \neq 0 \quad \textit{for } i = 1, \cdots, p, \quad j = 1, \cdots, q. \tag{1.3}$$

(iii) *The element* $f + \Psi$ *is a unit of the ring* $\mathbb{C}[x, y]/\Psi$.

(iv) *There exists a unique polynomial* $g \in \mathbb{C}[x, y]$ *such that*

$$fg \equiv 1 (\mathrm{mod}\ \Psi) \tag{1.4}$$

and the degree of g *in* x *is less than* p *and the degree in* y *is less than* q.

(v) *The equation*

$$\sum_{i=0}^{p-1}\sum_{k=0}^{q-1} f_{ik}F_a^iY(F_b^T)^k = (1, 0, \cdots, 0)_{1\times p}^T(1, 0, \cdots, 0)_{q\times 1} \tag{1.5}$$

is consistent.

Under condition (iii) above, the solution X can be expressed as $X = (f + \Psi)^{-1} * C$ and we have the following representation of X.

THEOREM 1.2. *If* (1.2) *is universally solvable and* $g(x, y) = \sum g_{rs}x^ry^s$ *is the polynomial given by* (1.4), *then*

$$X = \sum_{r=0}^{p-1}\sum_{s=0}^{q-1} g_{rs}A^rCB^s \tag{1.6}$$

is the solution of (1.2).

The polynomial g in (1.4) can be obtained from the auxiliary equation (1.5).

THEOREM 1.3. *Let* $Y = G \in \mathbb{C}^{p\times q}$ *be the solution of* (1.5). *Then*

$$g(x, y) = (1, x, \cdots, x^{p-1})G(1, y, \cdots, y^{p-1})^T$$

is the polynomial with the properties of (iv) *in Theorem* 1.1.

The representation of X in (1.6) yields a bound for the rank of X. For $L \in \mathbb{C}^{p\times r}$ and $R \in \mathbb{C}^{r\times q}$, put

$$K(A, L) = (L, AL, \cdots, A^{p-1}L) \tag{1.7}$$

and

$$D(B, R) = \begin{pmatrix} R \\ RB \\ \vdots \\ RB^{q-1} \end{pmatrix}.$$

THEOREM 1.4. *If* (1.2) *is universally solvable and if* C *is factorized as* $C = LR$, *then*

$$\operatorname{rank} X \leqq \min\{\operatorname{rank} K(A, L), \operatorname{rank} D(B, R)\}.$$

2. An algebraic approach to explicit solutions. We now consider our target equation

$$\sum_{i=0}^{p-1} A^i X D_i = C. \tag{1.1}$$

Whether (1.1) is universally solvable or not depends on the eigenvalues λ_i of A and on a polynomial matrix $D \in \mathbb{C}^{q \times q}[x]$,

$$D(x) = \sum_{i=0}^{p-1} D_i x^i.$$

THEOREM 2.1 (see [2], [4], or [14]). *Equation* (1.1) *is solvable for all* C *if and only if*

$$\det D(\lambda_i) \neq 0 \quad \textit{for } i = 1, \cdots, p. \tag{2.1}$$

The left-hand side of (1.1) will be regarded as a product of X and D. For that purpose, let us define for $F = \sum F_i x^i \in \mathbb{C}^{q \times q}[x]$ and for $M \in \mathbb{C}^{p \times q}$ the operation

$$M * F = \sum A^i M F_i. \tag{2.2}$$

Then $\mathbb{C}^{p \times q}$ becomes a right module over the ring $\mathbb{C}^{q \times q}[x]$. Recall that $a(x)$ is the characteristic polynomial of A and put

$$\Phi = a(x)\mathbb{C}^{q \times q}[x]$$

such that Φ is the ideal in $\mathbb{C}^{q \times q}[x]$ generated by $a(x)I$. By the Cayley–Hamilton theorem, we have $M * H = 0$ for all $H \in \Phi$. Hence it makes sense to define

$$M * (F + \Phi) = M * F, \tag{2.3}$$

and it is easy to verify that $\mathbb{C}^{p \times q}$ together with the operation (2.3) is a unitary right module over the ring $\mathbb{C}^{q \times q}[x]/\Phi$. For a nonzero polynomial matrix $H = \sum_{i=0}^{m} H_i x^i$ with $H_m \neq 0$ we put $\deg H = m$. In the case $H = 0$ we set $\deg H = -\infty$. Obviously, in each coset $F + \Phi$ there is a unique representative of degree less than p, $p = \deg a$.

Let us write (1.1) as

$$X * (D + \Phi) = C. \tag{2.4}$$

It is clear that (2.4) can be solved if $D + \Phi$ is invertible in $\mathbb{C}^{q \times q}[x]/\Phi$.

THEOREM 2.2. *The following statements* (i) *and* (ii) *are equivalent to condition* (2.1) *for the universal solvability of* (1.1):

(i) $D + \Phi$ *is a unit in the ring* $\mathbb{C}^{q \times q}[x]/\Phi$.

(ii) *There exists a unique matrix* $G \in \mathbb{C}^{q \times q}[x]$ *such that*

$$DG = GD \equiv I (\operatorname{mod} \Phi), \qquad \deg G < p. \tag{2.5}$$

Proof. We assume first that (2.1) holds. Then $d(x) = \det D$ and $a(x)$ have no zeros in common. Hence we have $a\gamma + d\delta = 1$ for some polynomials γ and δ. If $Q \in \mathbb{C}^{q \times q}[x]$ is the adjoint matrix of D, then $QD = DQ = dI$. Division of δQ by aI yields polynomial matrices B and G such that $\delta Q = aB + G$, $\deg G < \deg a = p$. From $\delta Q D = D \delta Q = I - a\gamma I$ we see that G satisfies (2.5). The condition $\det G < p$ together with (2.5) assures that the matrix G in (2.5) is uniquely determined. Obviously, (ii) implies (i). Finally, if

(i) holds, then (1.1) is solvable and

$$X = C * (D + \Phi)^{-1} \tag{2.6}$$

is the unique solution. □

From (2.6) we obtain the solution X in the following closed form.

THEOREM 2.3. *If* (1.1) *is universally solvable and*

$$G(x) = \sum_{i=0}^{p-1} G_i x^i \tag{2.7}$$

is the polynomial given by (2.5), *then*

$$X = \sum_{i=0}^{p-1} A^i C G_i \tag{2.8}$$

is the unique solution of (1.1).

The coefficient matrices G_i of (2.7) can be determined from an auxiliary equation which involves the companion matrix F_a of $a(x)$. For two matrices, $P = (p_{ij})$ and S, the Kronecker product is defined as $P \otimes S = (p_{ij}S)$. The identity matrix in (2.9) and (2.10) below is of size $q \times q$.

THEOREM 2.4. (i) *Let* $G = \sum_{i=0}^{p-1} G_i x^i$ *be given and put*

$$\tilde{G} = \begin{pmatrix} G_0 \\ G_1 \\ \vdots \\ G_{p-1} \end{pmatrix}, \quad E = \begin{pmatrix} I \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{2.9}$$

Then G *satisfies* $DG = GD \equiv I (\mathrm{mod}\, \Phi)$ *if and only if*

$$\sum_{i=0}^{p-1} (F_a \otimes I)^i \tilde{G} D_i = E \tag{2.10}$$

holds.

(ii) *The equation*

$$\sum_{i=0}^{p-1} (F_a \otimes I)^i Y D_i = E$$

is consistent if and only if (2.1) *holds.*

Proof. (i) Because of $\deg a = p$, there exists a unique matrix $U = (I, Ix, \cdots, Ix^{p-1})\tilde{U} \in \mathbb{C}^{q \times q}[x]$, which satisfies $GD \equiv U (\mathrm{mod}\, \Phi)$. We want to show that $\tilde{U}$ is the matrix on the left-hand side of (2.10). Put $\tilde{x} = (1, x, \cdots, x^{p-1})$. From $x^p = -a_0 - \cdots - a_{p-1}x^{p-1} + a(x)$, we obtain

$$\tilde{x}x = (x, \cdots, x^p) = \tilde{x}F_a + a(x)(0, \cdots, 0, 1)$$

and $\tilde{x}x^i = \tilde{x}F_a^i + av_i$ for some $v_i \in \mathbb{C}^{1 \times p}[x]$. Because of $G = (\tilde{x} \otimes I)\tilde{G}$, we have

$$GD = \sum_{i=0}^{p-1} (\tilde{x} \otimes I) x^i \tilde{G} D_i = \sum_{i=0}^{p-1} [(\tilde{x}F_a^i + av_i) \otimes I]\tilde{G}D_i$$

$$= (\tilde{x} \otimes I) \sum_{i=0}^{p-1} (F_a \otimes I)^i \tilde{G} D_i + a \sum_{i=0}^{p-1} (v_i \otimes I)\tilde{G}D_i .$$

Hence

$$\tilde{U} = \sum_{i=0}^{p-1} (F_a \otimes I)^i \tilde{G} D_i .$$

Clearly, $I = (\tilde{x} \otimes I)E$ and $U = (\tilde{x} \otimes I)\tilde{U}$. Therefore, $U = I$ is equivalent to (2.10). Theorem 2.4(ii) follows immediately from Theorem 2.2. □

We consider now a particular case where $D(x)$ has only one invariant factor (different from 1). Put $e = (1, 0, \cdots, 0)^T$, $e \in \mathbb{C}^p$.

THEOREM 2.5. *Assume that there exists a vector* $v \in \mathbb{C}^{1 \times q}$ *such that*

$$\operatorname{rank}\begin{pmatrix} D(z) \\ v \end{pmatrix} = q \quad \textit{for all } z \in \mathbb{C}. \tag{2.11}$$

Then the equation

$$\sum_{i=0}^{p-1} F_a^i Z D_i = ev \tag{2.12}$$

is consistent if and only if

$$\det D(\lambda_i) \neq 0 \tag{2.1'}$$

for all eigenvalues λ_i *of* F_a.

Proof. Clearly, (2.1) implies the solvability of (2.12). Suppose now that $\det D(\lambda) = 0$ for some eigenvalue λ of F_a and that (2.12) has a solution Z. Then there exists a vector $w = (w_1, \cdots, w_q)^T$, $w \neq 0$, such that $D(\lambda)w = 0$ and a vector $u = (u_1, \cdots, u_p)$, $u \neq 0$, such that

$$uF_a = \lambda u. \tag{2.13}$$

From

$$\begin{pmatrix} D(\lambda) \\ v \end{pmatrix} w = \begin{pmatrix} 0 \\ vw \end{pmatrix}$$

and (2.11) follows $vw \neq 0$, and (2.13) yields $u_1 \neq 0$. We multiply both sides of (2.12) by u and w. On the left-hand side we obtain

$$u\left(\sum F_a^i Z D_i\right) w = uZ\left(\sum \lambda^i D_i\right) w = uZD(\lambda) w = 0,$$

whereas on the right-hand side we have $u_1(vw) \neq 0$, which is a contradiction. Hence under the assumption (2.11) the condition (2.1) is also necessary for the consistency of (2.12). □

To estimate the rank of the solution X of (1.1), we use the representation (2.8).

THEOREM 2.6. *Assume that* (2.1) *holds and let* C *be factorized as* $C = LR$. *Then*

$$\operatorname{rank} X \leqq \min \{\operatorname{rank} K(A, L), \operatorname{rank} (I \otimes R)\tilde{G}\},$$

where $K(A, L)$ *is defined by* (1.7) *and* $\tilde{G}$ *is given by* (2.9).

Proof. Write X in (2.8) as

$$X = (L, AL, \cdots, A^{p-1}L)\begin{pmatrix} RG_0 \\ \vdots \\ RG_{p-1} \end{pmatrix} = K(A, L)(I \otimes R)\tilde{G}.$$

3. Realizations of D^{-1}. As its elements are rational functions, the matrix D^{-1} can be split into a proper rational and a polynomial part and thus be written as

$$D^{-1}(x) = K(Ix - F)^{-1}M + S(I - Nx)^{-1}T, \tag{3.1}$$

where N is nilpotent and the eigenvalues of F are zeros of det D (see, e.g., [9] or [6]). By a slight abuse of standard terminology, we will call a decomposition of the form (3.1) a *realization* of D^{-1}. Whenever a realization of D^{-1} is available, (1.1) can be reduced to a pair of equations of simpler type.

It will be convenient to express the solution of (1.1) by a contour integral.

THEOREM 3.1 [14]. *If the condition*

$$\det\left(\sum \lambda_\nu^i D^i\right) = \det D(\lambda_\nu) \neq 0, \qquad \nu = 1, \cdots, p \tag{2.1''}$$

holds, then the solution of (1.1) *is given by*

$$X = \frac{1}{2\pi i}\oint_\Gamma (Iz - A)^{-1} C D^{-1}(z)\,dz, \tag{3.2}$$

where Γ is a positively orientated simple closed curve such that all eigenvalues of A are interior to Γ and all zeros of det $D(z)$ *are exterior to* Γ.

We single out two special equations.

COROLLARY 3.2. *Let $\lambda_1, \cdots, \lambda_p$ and $\mu_1, \cdots, \mu_q$ be the eigenvalues of A and B, respectively.*

(i) *If $\lambda_i - \mu_j \neq 0$ for $i = 1, \cdots, p, j = 1, \cdots, q$, then*

$$AX - XB = C \tag{3.3}$$

has the unique solution

$$X = \frac{1}{2\pi i}\oint_\Gamma (Iz - A)^{-1} C (Iz - B)^{-1}\,dz,$$

where the λ_i's are in the interior of Γ and the μ_j's are outside of Γ.

(ii) *If $1 - \lambda_i\mu_j \neq 0$ for $i = 1, \cdots, p, j = 1, \cdots, q$, then*

$$X - AXB = C \tag{3.4}$$

has the unique solution

$$X = \frac{1}{2\pi i}\oint_\Gamma (Iz - A)^{-1} C (I - zB)^{-1}\,dz,$$

where all λ_i's are inside of Γ and for $\mu_j \neq 0$ all the numbers μ_j^{-1} are outside of Γ.

If we replace D^{-1} in the integrand of (3.2) by the realization (3.1) and apply Corollary 3.2, we are led to equations of type (3.3) and (3.4).

THEOREM 3.3. *Assume that* (2.1) *holds and let* (3.1) *be a realization of D^{-1}. Then the solution of* (1.1) *is given by*

$$X = X_1 M + X_2 T$$

where X_1 and X_2 are the solutions of

$$AX_1 - X_1 F = CK$$

and

$$X_2 - AX_2 N = CS,$$

respectively.

4. A more general equation with commuting matrices. Let $A_1, \cdots, A_n$ be complex $p \times p$ matrices which are pairwise commutative. In this section, we outline an algebraic approach for the equation

$$\sum A_1^{i_1}, \cdots, A_n^{i_n} X D_{i_1, \cdots, i_n} = C. \tag{4.1}$$

We assume that the i_ν's are nonnegative integers and that the sum in (4.1) is finite. The matrices $D_{i_1, \cdots, i_n}$ are of size $q \times q$ and $C \in \mathbb{C}^{p \times q}$.

Pairwise commutativity implies that the matrices $A_1, \cdots, A_n$ can be simultaneously triangularized [3]. Hence for some nonsingular S we have

$$S^{-1}AS = \operatorname{diag}(\lambda_{1i}, \cdots, \lambda_{pi}) + N_i, \qquad i = 1, \cdots, n,$$

where the matrices N_i are upper triangular and nilpotent. The vectors $\lambda_j = (\lambda_{j1}, \cdots, \lambda_{jn}), j = 1, \cdots, p$, are called *joint eigentuples* of $A_1, \cdots, A_n$. We associate (4.1) with a polynomial matrix $D \in \mathbb{C}^{q \times q}[y] = \mathbb{C}^{q \times q}[y_1, \cdots, y_n]$, where D is given by

$$D = \sum D_{i_1, \cdots, i_n} y_1^{i_1}, \cdots, y_n^{i_n}.$$

It is well known (see, e.g., [11]) that (4.1) is universally solvable if and only if

$$\det D(\lambda_j) \neq 0 \tag{4.2}$$

for all joint eigentuples λ_j of $A_1, \cdots, A_n$.

Let γ be the ideal of those polynomials in $\mathbb{C}[y]$ that vanish on $\{\lambda_1, \cdots, \lambda_p\}$,

$$\gamma = \bigcap_{j=1}^{p} (y_1 - \lambda_{j1}, \cdots, y_n - \lambda_{jn}),$$

and let $\varphi \subseteq \mathbb{C}[y]$ be the ideal given by

$$\varphi = \{f \mid f \in \mathbb{C}[y], f(A_1, \cdots, A_n) = 0\}.$$

Note that γ is determined by the semisimple part of the matrices A_i, whereas φ also involves the nilpotent parts. We have

$$\gamma^q \subseteq \varphi \subseteq \gamma,$$

which will allow us to work with γ^q instead of φ in what follows. We extend γ^q to an ideal Ψ in $\mathbb{C}^{q \times q}[y]$,

$$\Psi = \{hU \mid h \in \gamma^q, U \in \mathbb{C}^{q \times q}[y]\}.$$

For $F = \sum F_{k_1, \cdots, k_n} y_1^{k_1}, \cdots, y_n^{k_n} \in \mathbb{C}^{q \times q}[y]$ and $M \in \mathbb{C}^{p \times q}$, we define

$$M * (F + \Psi) = \sum A_1^{k_1}, \cdots, A_n^{k_n} M F_{k_1, \cdots, k_n}. \tag{4.3}$$

Then $\mathbb{C}^{p \times q}$ becomes a unitary right module over the ring $\mathbb{C}^{q \times q}[y]/\Psi$. Using Hilbert's Nullstellensatz, it can be shown that (4.2) holds if and only if $D + \Psi$ is a unit in $\mathbb{C}^{q \times q}[y]/\Psi$ or, equivalently, if there exists a matrix $G \in \mathbb{C}^{q \times q}[y]$ such that

$$DG = GD \equiv I (\operatorname{mod} \Psi). \tag{4.4}$$

If $G = \sum G_{k_1, \cdots, k_n} y_1^{k_1}, \cdots, y_n^{k_n}$ satisfies (4.4), then

$$X = \sum A_1^{k_1}, \cdots, A_n^{k_n} C G_{k_1, \cdots, k_n}$$

is the solution of (4.1).

REFERENCES

[1] T. E. Djaferis and S. K. Mitter, *Algebraic methods for the study of some linear matrix equations*, Linear Algebra Appl., 44 (1982), pp. 125–142.

[2] G. S. Datuashivili, *On the spectrum of a generalized matrix polynomial*, Bull. Acad. Sci. Georgian SSR, 44 (1966), pp. 7–9. (In Russian.)

[3] F. R. Gantmacher, *Matrizentheorie*, Springer-Verlag, Berlin, 1986.

[4] H. L. J. Hautus, *On the solvability of linear matrix equations*, Memorandum 1982-07, Department of Mathematics, Eindhoven University of Technology, the Netherlands, 1982.

[5] H. L. J. Hautus, *Linear matrix equations with applications to the regulator problem*, in Mathematical Tools and Models for Control, Systems Analysis and Signal Processing, Vol. 3 (Toulouse/Paris 1981/82), Centre National de la Recherche Scientifique, Paris, 1983, pp. 399–412.

[6] T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[7] R. E. Kalman, *On the Hermite–Fujiwara theorem in stability theory*, Quart. Appl. Math., 23 (1965), pp. 279–282.

[8] ———, *Algebraic characterization of polynomials whose zeroes lie in certain domains*, Proc. Nat. Acad. Sci. Math., 64 (1969), pp. 818–823.

[9] H. H. Rosenbrock, *Structural properties of linear dynamical systems*, Internat. J. Control, 20 (1974), pp. 191–202.

[10] J. J. Sylvester, *Sur la solution du cas le plus général des équations linéaires en quantités binaires c'est-à-dire en quaternions ou en matrices du second ordre*, C.R. Acad. Sci. Paris, 22 (1884), pp. 117–118.

[11] J. H. M. Wederburn, *Lectures on Matrices*, AMS Colloquium Publications XVII, American Mathematical Society, New York, 1934.

[12] H. K. Wimmer, *Linear matrix equations, controllability and observability, and the rank of solutions*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 570–578.

[13] ———, *Linear matrix equations: The module theoretic approach*, Linear Algebra Appl., 120 (1989), pp. 149–164.

[14] H. K. Wimmer and A. D. Ziebur, *Blockmatrizen und lineare Matrizengleichungen*, Math. Nachr., 59 (1974), pp. 213–219.

[15] N. J. Young, *Formulae for the solution of Lyapunov matrix equations*, Internat. J. Control, 31 (1980), pp. 159–179.