JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 79B. Nos. 3 and 4. July—December 1975

Similarity of Partitioned Matrices*

Robert B. Feinberg**

Institute for Basic Standards, National Bureau of Standards, Washington, D.C. 20234

(June 12, 1975)

Suppose that A, B, and T are matrices of order $r\times r$, $s\times s$, and $r\times s$ respectively over a field F. We prove that $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix}$ is similar to $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ iff $AX-XB=T$, for some matrix X. We also give some corollaries and a simple generalization.

Key words: Matric equation; partitioned matrix; rational canonical form; similarity.

Suppose that A, B, and T are matrices of order $r\times r$, $s\times s$, and $r\times s$ respectively over a commutative ring Φ. Let I_n denote the identity matrix of order n. If there is a matrix X of order $r\times s$ over Φ such that $AX-XB=T$, then it is a simple computation that

$$\begin{bmatrix} I_r & X \\ 0 & I_s \end{bmatrix}^{-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} I_r & X \\ 0 & I_s \end{bmatrix} = \begin{bmatrix} I_r & -X \\ 0 & I_s \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} I_r & X \\ 0 & I_s \end{bmatrix} = \begin{bmatrix} A & T \\ 0 & B \end{bmatrix}.$$

Thus $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix}$ is similar to $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ over Φ.

The main result in this paper (Theorem 6) is the converse to the above statement in the case when Φ is a field F, namely, if $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix}$ is similar to $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ over F, then there is a matrix X such that $AX-XB=T$. We also give some corollaries and a simple generalization of the theorem.

This result has been proven independently in [2],[1] and special cases of it have been established in [3], [4], and [6].

At this point we record some notation used throughout the paper. For integers r and s, let F_{rs} denote the collection of $r\times s$ matrices over F and let F'_{rr} denote the group of nonsingular matrices of order r. For $M, N\epsilon F_{rr}$, $M\tilde{S}N$ ($M\tilde{E}N$) represents the statement that M is similar (equivalent) to N over F. We denote the minimal polynomial of M by $f_M(x)$, and the companion matrix of $f_M(x)$ by $C(f_M(x))$. The rational canonical form of M is represented by $RF(M)$, and the minor obtained by deleting row i and column j is represented by $(M)_{ij}$. When the matrix M under discussion is understood, we let R_i denote the ith row of M and C_j denote the jth column. The elementary row operation of adding α times row j to row i is represented by $R_i\rightarrow R_i+\alpha R_j$.

See [5] for a good reference on matrices.

Let us note from the onset that in proving the main result we may assume w.l.o.g. (without loss of generality) that $A=RF(A)$ and $B=RF(B)$. Supposing that $U\epsilon F'_{rr}$ and $V'\epsilon F'_{ss}$ are such that $UAU^{-1}=RF(A)$ and $VAV^{-1}=RF(B)$, then

AMS Subject Classification: Primary 15A21, 15A24.

*This paper was prepared while the author was a National Academy of Sciences-National Research Council Postdoctoral Research Associate at the National Bureau of Standards, Washington, D.C. 20234. He would like to thank Dr. Morris Newman for several helpful comments.

**Present address: Department of Mathematics, Clarkson College of Technology, Potsdam, New York 13676.

[1] Figures in brackets indicate the literature references at the end of this paper.

$$\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \Longleftrightarrow \begin{bmatrix} RF(A) & UTV^{-1} \\ 0 & RF(B) \end{bmatrix} \tilde{S} \begin{bmatrix} RF(A) & 0 \\ 0 & RF(B) \end{bmatrix}$$

and

$$AX-XB=T \Longleftrightarrow RF(A)(UXV^{-1})-(UXV^{-1})RF(B)=UTV^{-1}.$$

Let us note also that we may assume w.l.o.g. that both A and B are nonzero. If both A and B are zero, the result is trivially true. If one of A and B is zero and the other is a multiple of the identity, the result is again trivially true. If A is zero and B is not a multiple of the identity, set $\tilde{A}=A+I=I$, $\tilde{B}=B+I$. Then both $\tilde{A}$ and $\tilde{B}$ are nonzero,

$$\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \Longleftrightarrow \begin{bmatrix} \tilde{A} & T \\ 0 & \tilde{B} \end{bmatrix} \tilde{S} \begin{bmatrix} A & 0 \\ 0 & \tilde{B} \end{bmatrix},$$

and $AX-XB=T \Longleftrightarrow \tilde{A}X-X\tilde{B}=T$. We obtain a similar result when the assumptions on A and B are interchanged. Hence in all cases we may assume w.l.o.g. that both A and B are nonzero.

It is now convenient to present the following well-known result. For an outline of the proof see [5, Ch. III, ex. 6 and 7].

LEMMA 0: *Suppose* $A \epsilon F_{rr}$, $B \epsilon F_{ss}$, *and that* A *and* B *have no eigenvalues in common. Then for all* $T \epsilon F_{rs}$ *there is an* $X \epsilon F_{rs}$ *such that* $T=AX-XB$.

Later on we will note a converse to this lemma.

We begin towards the proof of Theorem 6 by recording three technical lemmas. They contain essentially all the hard work.

LEMMA 1: *Suppose* $A_i \epsilon F_{u_i u_i}$, $1 \leq i \leq m$, $B_j \epsilon F_{v_j v_j}$, $1 \leq j \leq n$, $T_{i,r+j} \epsilon F_{u_i v_j}$, $1 \leq i \leq m$, $1 \leq j \leq n$.

Then

$$\left[\begin{array}{ccc|ccc} A_1 & & & T_{1,m+1} & \cdots & T_{1,m+n} \\ & \ddots & & \vdots & & \vdots \\ & & A_m & T_{m,m+1} & \cdots & T_{m,m+n} \\ \hline & & & B_1 & & \\ & 0 & & & \ddots & \\ & & & & & B_n \end{array}\right] \tilde{S} \left[\begin{array}{ccc|ccc} A_1 & & & & & \\ & \ddots & & & 0 & \\ & & A_m & & & \\ \hline & & & B_1 & & \\ & 0 & & & \ddots & \\ & & & & & B_n \end{array}\right]$$

$\Rightarrow$ for $1 \leq i \leq m$, $1 \leq j \leq n$,

$$\begin{bmatrix} A_i & T_{i,m+j} \\ 0 & B_j \end{bmatrix} \tilde{S} \begin{bmatrix} A_i & 0 \\ 0 & B_j \end{bmatrix}.$$

PROOF: Let

$$M=\left[\begin{array}{ccc|ccc} \lambda I-A_1 & & & -T_{1,m+1} & \cdots & -T_{1,m+n} \\ & \ddots & & \vdots & & \vdots \\ & & \lambda I-A_m & -T_{m,m+1} & \cdots & -T_{m,m+n} \\ \hline & & & \lambda I-B_1 & & \\ & 0 & & & \ddots & \\ & & & & & \lambda I-B_n \end{array}\right]$$

and $D=\text{diag}[\lambda I-A_1, \ldots, \lambda I-A_m, \lambda I-B_1, \ldots, \lambda I-B_n]$, so that M and D are matrices over the principal ideal domain $F[\lambda]$. The hypotheses, together with the fundamental theorem on similarity over a field, imply that M and D are equivalent over $F[\lambda]$.

Now let M_1 be obtained from M by replacing $T_{1,m+1}, \ldots, T_{1,m+n}$ with blocks of 0's. Note that to obtain a minor of M_1 with nonzero determinant, it is necessary that the number of rows deleted which pass through the block $\lambda I - A_1$, equal the number of columns deleted which pass through this block. It follows from this that every determinantal minor of M_1 is a determinantal minor of M also. Since $M\tilde{E}D$, we obtain that $M_1\tilde{E}D$ as well. Write M_1 as $(\lambda I - A_1) \dotplus M_2$ and D as $(\lambda I - A_1) \dotplus D_2$. It then follows from [5, Ch. 2, ex. 1] that $M_2\tilde{E}D_2$. Repeating this process m times, we obtain that $M_m\tilde{E}D_m$, where

$$M_m = \left[\begin{array}{c|cccc} \lambda I - A_m & -T_{m,m+1} & \cdots & & -T_{m,m+n} \\ \hline & \lambda I - B_1 & & & \cdot \\ & & \cdot & & \cdot \\ 0 & & & \cdot & \cdot \\ & & & & \lambda I - B_n \end{array}\right]$$

and $D = \operatorname{diag}[\lambda I - A_m, \lambda I - B_1, \ldots, \lambda I - B_n]$.

Arguing analogously on the columns of M_m, we obtain finally that

$$\left[\begin{array}{c|c} \lambda I - A_m & -T_{m,m+1} \\ \hline 0 & \lambda I - B_1 \end{array}\right] \tilde{E} \left[\begin{array}{c|c} \lambda I - A_m & 0 \\ \hline 0 & \lambda I - B_1 \end{array}\right]$$

from which it follows that

$$\left[\begin{array}{c|c} A_m & T_{m,m+1} \\ \hline 0 & B_1 \end{array}\right] \tilde{S} \left[\begin{array}{c|c} A_m & 0 \\ \hline 0 & B_1 \end{array}\right].$$

This establishes the lemma in the case when $i = m$ and $j = 1$.

To prove the lemma for arbitrary $(i, j) \epsilon [1, m] \times [1, n]$, note that by simultaneous row and column permutation we may obtain

$$\left[\begin{array}{cccccc|cccccc} A_1 & & & & & & \cdot & \cdots & & & & \cdot \\ & \ddots & & & & & \vdots & & & & & \vdots \\ & & \hat{A}_i & & & & & & & & & \\ & & & \ddots & & & \vdots & & & & & \vdots \\ & & & & A_m & & & & & & & \\ & & & & & A_i & T_{i,m+j} & \cdots & & & & \cdot \\ \hline & & & & & & B_j & & & & & \\ & & & & & & & B_1 & & & & \\ & & 0 & & & & & & \ddots & & & \\ & & & & & & & & & \hat{B}_j & & \\ & & & & & & & & & & \ddots & \\ & & & & & & & & & & & B_n \end{array}\right] \tilde{S} \left[\begin{array}{cccccc|cccccc} A_1 & & & & & & & & & & & \\ & \ddots & & & & & & & & & & \\ & & \hat{A}_i & & & & & & & 0 & & \\ & & & \ddots & & & & & & & & \\ & & & & A_m & & & & & & & \\ & & & & & A_i & & & & & & \\ \hline & & & & & & B_j & & & & & \\ & & & & & & & B_1 & & & & \\ & & 0 & & & & & & \ddots & & & \\ & & & & & & & & & \hat{B}_j & & \\ & & & & & & & & & & \ddots & \\ & & & & & & & & & & & B_n \end{array}\right]$$

Running the above argument on this new pair of matrices, we get finally that

$$\left[\begin{array}{c|c} A_i & T_{i,m+j} \\ \hline 0 & B_j \end{array}\right] \tilde{S} \left[\begin{array}{c|c} A_i & 0 \\ \hline 0 & B_j \end{array}\right].$$

Q.E.D.

LEMMA 2: *Suppose* $A \epsilon F_{rr}$, $B \epsilon F_{ss}$, *and* $T \epsilon F_{rs}$, *where both* A *and* B *are nonderogatory and in rational canonical form. Then (a)* $\exists \bar{X}, \bar{T} \epsilon F_{rs}$ *such that* $\bar{T}$ *has nonzero entries only in its first column and* $T - \bar{T} = A\bar{X} - \bar{X}B$. *Also,* $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & \bar{T} \\ 0 & B \end{bmatrix}$. *(b)* $\exists \bar{\bar{X}}, \bar{\bar{T}} \epsilon F_{rs}$ *such that* $\bar{\bar{T}}$ *has nonzero entries only in its last row and* $T - \bar{\bar{T}} = A\bar{\bar{X}} - \bar{\bar{X}}B$. *Also,* $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & \bar{\bar{T}} \\ 0 & B \end{bmatrix}$.

PROOF: By assumption,

$$A = \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 0 & 0 & 0 & \dots & 1 \\ -\gamma_0 & -\lambda_1 & & \dots & -\gamma_{r-1} \end{bmatrix} = C(f_A(x))$$

and

$$B = \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 0 & 0 & 0 & \dots & 1 \\ -\eta_0 & -\eta_1 & & & -\eta_{s-1} \end{bmatrix} = C(f_B(x)),$$

where

$$f_A(x) = \gamma_0 + \gamma_1 x + \dots + \gamma_{r-1}x^{r-1} + x^r$$

and

$$f_B(x) = \eta_0 + \eta_1 x + \dots + \eta_{s-1}s^{s-1} + x^s.$$

Write $T = (t_{i,r+j})_{\substack{1 \leq i \leq r \\ 1 \leq j \leq s}}$.

(a) Perform the following $s-1$ sequences of elementary row and column operations on $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix}$, obtaining a sequence of matrices

$$\left\{ \begin{bmatrix} A & T_j \\ 0 & B \end{bmatrix} \right\}_{j=1}^{s-1}:$$

Sequence 1: For $1 \leq i \leq r$,

$$R_i \rightarrow R_i - t_{i,r+s} R_{r+s-1}$$

$$C_{r+s-1} \rightarrow C_{r+s-1} + t_{i,r+s} C_i.$$

These operations are effected by the similarity

$$\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \rightarrow \begin{bmatrix} I_r & X_1 \\ 0 & I_s \end{bmatrix} \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \begin{bmatrix} I_r & -X_1 \\ 0 & I_s \end{bmatrix} \equiv \begin{bmatrix} A & T_1 \\ 0 & B \end{bmatrix},$$

for an appropriate $X_1 \epsilon F_{rs}$. Note that the last column of $T_1 \equiv (t^1_{k,r+l})_{\substack{1 \leq k \leq r \\ 1 \leq l \leq s}}$ consists entirely of 0's.

Sequence j $(2 \leq j \leq s-1)$: For $1 \leq i \leq r$,

$$R_i \to R_i - t^{j-1}_{i,r+s-(j-1)} R_{r+s-j}$$

$$C_{r+s-j} \to C_{r+s-j} + t^{j-1}_{i,r+s-(j-1)} C_i.$$

These operations are effected by the similarity

$$\begin{bmatrix} A & T_{j-1} \\ 0 & B \end{bmatrix} \to \begin{bmatrix} I_r & X_j \\ 0 & I_s \end{bmatrix} \begin{bmatrix} A & T_{j-1} \\ 0 & B \end{bmatrix} \begin{bmatrix} I_r & -X_j \\ 0 & I_s \end{bmatrix} \equiv \begin{bmatrix} A & T_j \\ 0 & B \end{bmatrix},$$

for an appropriate $X_j \epsilon F_{rs}$. Note that the last j columns of $T_j \equiv (t^j_{k,r+l})_{\substack{1 \leq k \leq r \\ 1 \leq l \leq s}}$ consist entirely of 0's.

Now let $\bar{X} = \sum_{j=1}^{s-1} X_j$ and let $\bar{T} = T_{s-1}$. Then $\bar{T}$ has nonzero entries only in its first column and

$$\begin{bmatrix} I_r & \bar{X} \\ 0 & I_s \end{bmatrix} \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \begin{bmatrix} I_r & -\bar{X} \\ 0 & I_s \end{bmatrix}$$

$$= \left\{ \prod_{j=1}^{s-1} \begin{bmatrix} I_r & X_{s-j} \\ 0 & I_s \end{bmatrix} \right\} \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \left\{ \prod_{j=1}^{s-1} \begin{bmatrix} I_r & -X_j \\ 0 & I_s \end{bmatrix} \right\} = \begin{bmatrix} A & \bar{T} \\ 0 & B \end{bmatrix}$$

It follows from this that $T - \bar{T} = A\bar{X} - \bar{X}B$, also that

$$\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & \bar{T} \\ 0 & B \end{bmatrix}.$$

(b) Perform the following $r-1$ sequences of elementary row and column operations on $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix}$, obtaining a sequence of matrices $\left\{ \begin{bmatrix} A & U_i \\ 0 & B \end{bmatrix} \right\}_{i=1}^{r-1}$:

Sequence 1. For $1 \leq j \leq s$,

$$C_{r+j} \to C_{r+j} - t_{1,r+j} C_2$$

$$R_2 \to R_2 + t_{1,r+j} R_{r+j}.$$

These operations are effected by the similarity

$$\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \to \begin{bmatrix} I_r & Y_1 \\ 0 & I_s \end{bmatrix} \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \begin{bmatrix} I_r & -Y_1 \\ 0 & I_s \end{bmatrix} \equiv \begin{bmatrix} A & U_1 \\ 0 & B \end{bmatrix},$$

for an appropriate $Y_1 \epsilon F_{rs}$. Note that the first row of $U_1 \equiv (u^1_{k,r+1})_{\substack{1 \leq k \leq r \\ 1 \leq l \leq s}}$ consists entirely of 0's.

Sequence i $(2 \leq i \leq r-1)$. For $1 \leq j \leq s$,

$$C_{r+j} \to C_{r+j} - u^{i-1}_{i,r+j} C_{i+1}$$

$$R_{i+1} \to R_{i+1} + u^{i-1}_{i,r+j} R_{r+j}.$$

These operations are effected by the similarity

$$\begin{bmatrix} A & U_{i-1} \\ 0 & B \end{bmatrix} \rightarrow \begin{bmatrix} I_r & Y_i \\ 0 & I_s \end{bmatrix} \begin{bmatrix} A & U_{i-1} \\ 0 & B \end{bmatrix} \begin{bmatrix} I_r & -Y_i \\ 0 & I_s \end{bmatrix} \equiv \begin{bmatrix} A & U_i \\ 0 & B \end{bmatrix},$$

for an appropriate $Y_i \epsilon F_{rs}$. Note that the first i rows of $U_i \equiv (u^i_{k,r+l})_{\substack{1 \leq k \leq r \\ 1 \leq l \leq s}}$ consist entirely of 0's.

Now let $\bar{\bar{X}} = \sum_{i=1}^{r-1} Y_i$ and let $\bar{\bar{T}} = U_{r-1}$. Then $\bar{\bar{T}}$ has nonzero entries only in its last row and

$$\begin{bmatrix} I_r & \bar{\bar{X}} \\ 0 & I_s \end{bmatrix} \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \begin{bmatrix} I_r & -\bar{\bar{X}} \\ 0 & I_s \end{bmatrix} = \left\{ \prod_{i=1}^{r-1} \begin{bmatrix} I_r & U_{r-i} \\ 0 & I_s \end{bmatrix} \right\} \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \left\{ \prod_{i=1}^{r-1} \begin{bmatrix} I_r & -U_i \\ 0 & I_s \end{bmatrix} \right\} = \begin{bmatrix} A & \bar{\bar{T}} \\ 0 & B \end{bmatrix}.$$

It follows from this that $T - \bar{\bar{T}} = A\bar{\bar{X}} - \bar{\bar{X}}B$, also, that

$$\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & \bar{\bar{T}} \\ 0 & B \end{bmatrix}.$$

Q.E.D.

LEMMA 3: *Suppose* $A \epsilon F_{rr}$, $B \epsilon F_{ss}$, *and* $T \epsilon F_{rs}$, *where both* A *and* B *are nonderogatory and in rational canonical form. Assume also that* f_A *and* f_B *are both powers of the same monic irreducible polynomial* p(x) *and that* $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$. *Then* T *must be* 0 *if*

(*a*) $r \leq s$ *and* T *has nonzero entries possibly only in its first column*

or

(*b*) $s \leq r$ *and* T *has nonzero entries possibly only in its last row.*

PROOF: By hypothesis,

$$A = \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 0 & 0 & 0 & \ldots & 1 \\ -\gamma_0 & -\gamma_1 & & & -\gamma_{r-1} \end{bmatrix} = C(f_A)$$

and

$$B = \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 0 & 0 & 0 & \ldots & 1 \\ -\eta_0 & -\eta_1 & & \ldots & -\eta_{s-1} \end{bmatrix} = C(f_B)$$

where

$$f_A(x) = \gamma_0 + \gamma_1 x + \ldots + \gamma_{r-1}x^{r-1} + x^r = p(x)^e$$

and

$$f_B(x) = \eta_0 + \eta_1 x + \ldots + \eta_{s-1}x^{s-1} + x^s = p(x)^f,$$

for some integers e and f. Write $T = (t_{i,r+j})_{\substack{1 \leq i \leq r \\ 1 \leq j \leq s}}$.

Now let $D=\text{diag}[\lambda I-A,\ \lambda I-B]\in F[\lambda]_{(r+s),(r+s)}$. For $(D)_{ij}$ any $(r+s-1)\times(r+s-1)$ minor of D, it may be seen that det $(D)_{ij}\neq 0 \Rightarrow i, j\leqslant r$ or $i, j\geqslant r+1$. Note also that for $i, j\leqslant r$, det $(D)_{ij}=g_{ij}(\lambda)f_B(\lambda)$, for some $g_{ij}(\lambda)\in F[\lambda]$, where $gr_i(\lambda)=\pm 1$; and that for $i, j\geqslant r+1$, det $(D)_{ij}=h_{ij}(\lambda)f_A(\lambda)$, for some $h_{ij}(\lambda)\in F[\lambda]$, where $h_{r+s,\,r+1}(\lambda)=\pm 1$.

Let Δ_{r+s-1} be the $(r+s-1)\times(r+s-1)$determinantal divisor of D. It then follows from the above calculations that

$$\Delta_{r+s-1}=g.c.d.\ \{f_A(\lambda), f_B(\lambda)\} = \begin{cases} f_A(\lambda), & \text{if } r\leqslant s \\ f_B(\lambda), & \text{if } s\leqslant r. \end{cases}$$

Note also that $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \Rightarrow \begin{bmatrix} \lambda I-A & -T \\ 0 & \lambda I-B \end{bmatrix}$ is equivalent over $F[\lambda]$ to D. Hence for all $i, j\geqslant r+s$,

$$\Delta_{r+s-1} \mid \det\left(\begin{bmatrix} \lambda I-A & -T \\ 0 & \lambda I-B \end{bmatrix}_{ij}\right).$$

We now prove (a) and (b).

(a) It may be seen that for $i\leqslant r$, $\det\left(\begin{bmatrix} \lambda I-A & -T \\ 0 & \lambda I-B \end{bmatrix}\right)_{r+s,\,i} = \pm t_{i,\,r+1}\ \lambda^{r-1}+q(\lambda)+E$, where $q(\lambda)$ is a polynomial of degree $\leqslant r-2$ and E is an $(r+s-1)\times(r+s-1)$ determinantal minor of D. It follows that

$$\Delta_{r+s-1} | \pm t_{i,\,r+1}\ \lambda^{r-1}+q(\lambda)+E.$$

Since $\Delta_{r+s-1}|E$ and since $\Delta_{r+s-1}=f_A(\lambda)$ is a polynomial of degree r in this case, we obtain that $t_{i,\,r+1}=0$, $i\leqslant r$, whence $T=0$.

(b) It may be seen that $\det\left(\begin{bmatrix} \lambda I-A & -T \\ 0 & \lambda I-B \end{bmatrix}\right)_{r+s,\,1} = \pm t_{r,\,r+1} \pm t_{r,\,r+2}\ \lambda \pm t_{r,\,r+3}\ \lambda^2 \pm \ .\ .\ .$ $\pm t_{r,r+s}\ \lambda^{s-1}+F$, where F is an $(r+s-1)\times(r+s-1)$ determinantal minor of D. It follows that

$$\Delta_{r+s-1} | \pm t_{r,\,r+1} \pm t_{r,\,r+2}\ \lambda \pm\ .\ .\ .\ \pm t_{r,\,r+s}\ \lambda^{s-1}+F.$$

Since $\Delta_{r+s-1}|F$ and since $\Delta_{r+s-1}=f_B(\lambda)$ is a polynomial of degree s in this case, we obtain that $t_{r,\,r+1}=t_{r,\,r+2}=\ .\ .\ .\ =t_{r,\,r+s}=0$, whence $T=0$. Q.E.D.

It is now convenient to prove our main result in a simple special case.

Lemma 4: Suppose $A\in F_{rr}$, $B\in F_{ss}$, *and* $T\in F_{rs}$, *where both* A *and* B *are nonderogatory. Assume also that* $f_A(x)=p_1(x)^d$ *and* $f_B(x)=p_2|x(^e$, *where* $p_1(x)$ *and* $p_2(x)$ *are monic irreducible polynomials.*

Then $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \Rightarrow T=AX-XB$, *for some* $X\in F_{rs}$.

PROOF: As noted above, we may assume w.l.o.g. that $A=RF(A)$ and $B=RF(B)$. If $p_1(x)\neq p_2(x)$, then A and B have no eigenvalues in common, and hence we know from Lemma 0 that $\exists X\in F_{rs}$ such that $AX-XB=T$ (the hypothesis on similarity is superfluous in this case.)

Assume now that $p_1(x)=p_2(x)$. If $r\leqslant s$, use Lemma 2a to find $\bar{X}, \bar{T}\in F_{rs}$ such that $\bar{T}$ has nonzero entries only in its first column, $T-\bar{T}=A\bar{X}-\bar{X}B$, and $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & T \\ 0 & B \end{bmatrix}$. Since $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, it follows that $\begin{bmatrix} A & \bar{T} \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, and we then obtain from Lemma 3a that $\bar{T}=0$. Thus $T=A\bar{X}-\bar{X}B$ in this case. If $s \leqslant r$, use Lemma 2b to find $\bar{\bar{X}}$, $\bar{\bar{T}} \epsilon F_{rs}$ such that $\bar{\bar{T}}$ has nonzero entries only in its last row, $T-\bar{\bar{T}}=A\bar{\bar{X}}-\bar{\bar{X}}B$, and $\begin{bmatrix} A & \bar{\bar{T}} \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & T \\ 0 & B \end{bmatrix}$. Again, $\begin{bmatrix} A & \bar{\bar{T}} \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, and we then obtain from Lemma 3b that $\bar{\bar{T}}=0$. Thus $T=A\bar{\bar{X}}-\bar{\bar{X}}B$ in this case as well.

Q.E.D.

We now drop the requirement that A and B be nonderogatory.

LEMMA 5: *Suppose* $A \epsilon F_{rr}$, $B \epsilon F_{ss}$, *and* $T \epsilon F_{rs}$. *Assume that both* A *and* B *are nonzero and that* $f_A(x)=p_1(x)^d$ *and* $f_B(x)=p_2(x)^e$, *where* $p_1(x)$ *and* $p_2(x)$ *are monic irreducible polynomials. Then* $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \Rightarrow T=AX-XB$ *for some* $X \epsilon F_{rs}$.

PROOF: As before, we may assume that $A=RF(A)$, $B=RF(B)$, and that $p_1(x)=p_2(x)$. We then have that $A=\text{diag}[C(p(x)^{d_1}), C(p(x)^{d_2}), \ldots, C(p(x)^{d_u})]$, where $d=d_1 \geqslant d_2 \geqslant \ldots \geqslant d_u$, and $B=\text{diag}[C(p(x)^{e_1}), C(p(x)^{e_2}), \ldots, C(p(x)^{e_v})]$, where $e=e_1 \geqslant e_2 \geqslant \ldots \geqslant e_v$. Now write $T=(T_{i,u+j})_{\substack{1 \leqslant i \leqslant u \\ 1 \leqslant j \leqslant v}}$, where $T_{i,u+j}$ has d_i rows and e_j columns. We then have by Lemma 1 that for all $(i, j) \epsilon [1, u] \times [1, v]$,

$$\begin{bmatrix} C(p(x)^{d_i}) & T_{i,u+j} \\ 0 & C(p(x)^{e_j}) \end{bmatrix}$$

$$\tilde{S} \begin{bmatrix} C(p(x)^{d_i}) & 0 \\ 0 & C(p(x)^{e_j}) \end{bmatrix}.$$

It then follows from Lemma 4 that there is a matrix $X_{i,u+j}$ over F such that

$$T_{i,u+j}=C(p(x)^{d_i})\, X_{i,u+j}-X_{i,u+j}\, C(p(x)^{e_j}).$$

Let $X=(X_{i,u+j})_{\substack{1 \leqslant i \leqslant u \\ 1 \leqslant j \leqslant v}} \epsilon F_{rs}$. We then obtain by straightforward computation that $T=AX-XB$.

Q.E.D.

We now establish the main result.

THEOREM 6: *Suppose* $A \epsilon F_{rr}$, $B \epsilon F_{ss}$, *and* $T \epsilon F_{rs}$. *Then* $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \Rightarrow T=AX-XB$, *for some* $X \epsilon F_{rs}$.

PROOF: As before, we may assume that both A and B are nonzero and in rational canonical form. Assume also that

$$f_A(x)=p_1(x)^{d_1}p_2(x)^{d_2} \ldots p_u(x)^{d_u}$$

and

$$f_B(x)=q_1(x)^{e_1}q_2(x)^{e_2} \ldots q_v(x)^{e_v},$$

where $\{p_i(x)\}_{i=1}^{u}$ and $\{q_j(x)\}_{j=1}^{v}$ are sets of distinct irreducible polynomials in $F[x]$ We may then write $A=\text{diag}[G_1, \ldots, G_u]$ and $B=\text{diag}[H_1, \ldots, H_v]$, where $f_{G_i}(x)=p_1(x)^{d_i}$ and $f_{H_j}(x)=q_j(x)^{e_j}$. Now write $T=(T_{i,u+j})_{\substack{1 \leqslant i \leqslant u \\ 1 \leqslant j \leqslant v}}$, where $T_{i,u+j}$ is conformable with G_i and H_j. We then have by Lemma 1 that for all $(i, j) \epsilon [1, u] \times [1, v]$, $\begin{bmatrix} G_i & T_{i,u+j} \\ 0 & H_j \end{bmatrix} \tilde{S} \begin{bmatrix} G_i & 0 \\ 0 & H_j \end{bmatrix}$. It then follows from Lemma 5 that there is a matrix $X_{i,u+j}$ over F such that $T_{i,u+j}=G_iX_{i,u+j}-X_{i,u+j}H_j$. Let $X=(X_{i,u+j})_{\substack{1 \leqslant i \leqslant u \\ 1 \leqslant j \leqslant v}} \epsilon F_{rs}$. We then obtain by straightforward computation that $T=AX-XB$. Q.E.D.

COROLLARY 6.1: *Suppose* $A \epsilon F_{rr}$, $B \epsilon F_{ss}$, *and* T, $\tilde{T} \epsilon F_{rs}$. *Then*

$$\begin{bmatrix} A & T-\tilde{T} \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \Rightarrow \begin{bmatrix} A & \tilde{T} \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & \tilde{T} \\ 0 & B \end{bmatrix}.$$

PROOF: The hypotheses, together with Theorem 6, imply $T-\tilde{T}=AX-XB$, for some $X \epsilon F_{rs}$. It is then a simple computation that

$$\begin{bmatrix} I_r & -X \\ 0 & I_s \end{bmatrix} \begin{bmatrix} A & \tilde{T} \\ 0 & B \end{bmatrix} \begin{bmatrix} I_r & X \\ 0 & I_s \end{bmatrix} = \begin{bmatrix} A & T \\ 0 & B \end{bmatrix},$$

so that

$$\begin{bmatrix} A & \tilde{T} \\ 0 & B \end{bmatrix} \tilde{S} \begin{bmatrix} A & T \\ 0 & B \end{bmatrix}.$$

Q.E.D.

Note that the converse to Corollary 6.1 fails. For example, let F=the reals, R, $A=B=(3)$, $T=(4)$, and $\tilde{T}=(2)$. Then $\begin{bmatrix} 3 & 4 \\ 0 & 3 \end{bmatrix} \tilde{S} \begin{bmatrix} 3 & 2 \\ 0 & 3 \end{bmatrix}$, since $\begin{bmatrix} 3-x & 4 \\ 0 & 3-x \end{bmatrix}$ is equivalent over $R[x]$ to $\begin{bmatrix} 3-x & 2 \\ 0 & 2 \end{bmatrix}$, but $\begin{bmatrix} 3 & 4-2 \\ 0 & 3 \end{bmatrix}$ is *not* similar to $\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$, since there is obviously no x satisfying $4-2=x(3-3)$.

We note two further consequences of the results and techniques developed thus far. First, they may be used to prove the converse to Lemma 0, namely that if $A \epsilon F_{rr}$ and $B \epsilon F_{ss}$ have the property that for all $T \epsilon F_{rs}$ there is an $X \epsilon F_{rs}$ such that $T=AX-XB$, then A and B have no eigenvalues in common. Second, they may be used to find an explicit solution in X of the matric equation $T=AX-XB$, at least in the case when A and B are in rational cononical form. See [1] for another approach to solving this equation.

We conclude with a simple generalization of Theorem 6.

THEOREM 7: *Suppose* $U_i \epsilon F_{r_i r_i}$, $1 \leq i \leq k$, *and* $N_{ij} \epsilon F_{r_i r_j}$, $1 \leq i < j \leq k$. *Then*

$$\begin{bmatrix} U_1 & & N_{ij} \\ & \ddots & \\ & & U_k \end{bmatrix} \tilde{S} \begin{bmatrix} U_1 & & \\ & \ddots & \\ & & U_k \end{bmatrix} \Leftrightarrow$$

for each $i, j \leq k \; \exists X_{ij} \epsilon F_{r_i r_j}$ *such that* $U_i X_{ij} - X_{ij} U_j = N_{ij}$.

References

[1] Bartels, R., and Stewart, G. W., Solution of the matrix equation $AX+XB=C$, Comm. ACM **15**, 820–827 (1972).
[2] Gustafson, W., On the matrix equations $AX-YB=C$ and $AX-XB=C$.
[3] Johnson, C. R., and Newman, M., A Condition for the diagonalizability of a partitioned matrix, J. Res. Nat. Bur. Stand. (U.S.) 79B, (Math. Sci.) Nos. 3 & 4, (1975).
[4] Kucera, V., The matrix equation $AX+XB=C$, Siam J. Appl. Math. **26**, 15–25 (1974).
[5] Newman, M., Integral matrices (Academic Press, New York, 1972).
[6] Roth, W., The equations $AX-YB=C$ and $AX-XB=C$ in Matrices, Proc. AMS **3**, 392–396, (1952).

(Paper 79B3 & 4–427)