JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 80B, No. 1, January–March 1976

Equivalence of Partitioned Matrices*

Robert B. Feinberg**

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

(November 18, 1975)

It is shown that if $M=\begin{bmatrix} A & T \\ 0 & B \end{bmatrix}$ is a partitioned matrix over a principal ideal domain R such that the matrices A and B are both square, then M is equivalent to $A\dotplus B$ $\langle=\rangle$ the matric equation $T=AY+XB$ is solvable. The result is generalized to treat the case when

$$M=\begin{bmatrix} M_{11} & M_{12} & \dots & M_{1t} \\ 0 & M_{22} & \dots & . \\ . & & & . \\ . & & & . \\ 0 & \dots & 0 & M_{tt} \end{bmatrix},$$

where each M_{ii} is square.

Key words: Determinantal divisors; equivalence; matric equation; partitioned matrix; Smith normal form.

Let R be a principal ideal domain and let R_{mn} denote the collection of $m \times n$ matrices over R. According to Theorem 2 of [2], if $A\epsilon R_{rr}$, $B\epsilon R_{ss}$, and $(\det A, \det B) = 1$, then for any $T\epsilon R_{rs}$,

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

where $S(M)$ denotes the Smith normal form of a matrix M. The proof consists essentially of establishing two elementary propositions:

(i) For arbitrary $A\epsilon R_{rr}$, $B\epsilon R_{ss}$, and $T\epsilon R_{rs}$, if the matric equation (*) $T = AY + XB$ has a solution X, $Y\epsilon R_{rs}$, then

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

where $\tilde{E}$ denotes equivalence of matrices (ii). In the case when $(\det A, \det B) = 1$, (*) is always solvable.

The central result of this note (Theorem 1) provides a converse to (i), namely that if

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

*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 conversations.

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

then (*) must be solvable. We generalize this result to the case when

$$M=\begin{bmatrix} M_{11} & M_{12} & \dots & M_{1t} \\ 0 & M_{22} & \dots & \\ \vdots & & & \vdots \\ 0 & 0 & & M_{tt} \end{bmatrix},$$

where each M_{ii} is square, and also derive some corollaries.

Subsequent to completion of this work, the author discovered that Theorem 1 had been established in [3] in the case when R is the domain of polynomials over a field. The proof there carries over immediately to the case when R is an arbitrary P.I.D., and is similar to the proof of Theorem 1 presented here. The generalization of Theorem 1 is not developed there, however.

In the sequal R'_n will denote the group of unimodular $n \times n$ matrices over R, I_n will denote the identity matrix of order n, I will denote an identity matrix of unspecified order, 0_{mn} will denote the 0 matrix of order $m \times n$, 0_m will denote 0_{mm}, and $d_k[M]$ will denote the kth determinantal divisor of the matrix M.

See [1] for a good general reference on matrices over a P.I.D.

THEOREM 1: *Let* R *be a P.I.D.,* $A \epsilon R_{rr}$, $B \epsilon R_{ss}$, *and* $T \epsilon R_{rs}$. *Then* $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \langle = \rangle T = AY + XB$, *for suitable* $X, Y \epsilon R_{rs}$.

PROOF: $(\langle =)$ Note that $\begin{bmatrix} I_r & X \\ 0 & I_s \end{bmatrix}, \begin{bmatrix} I_r & Y \\ 0 & I_s \end{bmatrix} \epsilon R'_{r+s}$, and that

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

Hence

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

$(= \rangle)$ Let $\varphi \begin{bmatrix} A & T \\ 0 & B \end{bmatrix}$ be the statement we wish to prove, namely $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \langle = \rangle \exists X, Y \epsilon R_{r+s}$ such that $T = AY + XB$.

We will begin with four reduction steps

(i) We may assume w.l.o.g. (without loss of generality) that $A = S(A)$, $B = S(B)$. Justification: Choose $U, U^* \epsilon R'_{rr}$; $V, V^* \epsilon R'_{ss}$ such that $UAU^* = S(A)$, $VAV^* = S(B)$. Note that

$$\begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} U^* & 0 \\ 0 & V^* \end{bmatrix} = \begin{bmatrix} S(A) & 0 \\ 0 & S(B) \end{bmatrix}$$

and that

$$\begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \begin{bmatrix} U^* & 0 \\ 0 & V^* \end{bmatrix} = \begin{bmatrix} S(A) & \tilde{T} \\ 0 & S(B) \end{bmatrix},$$

where $\tilde{T} = UAV^*$.

Hence $$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} S(A) & 0 \\ 0 & S(B) \end{bmatrix} \text{ and } \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} S(A) & \tilde{T} \\ 0 & S(B) \end{bmatrix}.$$

Thus $$\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} <=> \begin{bmatrix} S(A) & \tilde{T} \\ 0 & S(B) \end{bmatrix} \tilde{E} \begin{bmatrix} S(A) & 0 \\ 0 & S(B) \end{bmatrix}.$$

Note also that $T = AY + XB <=> UTV^* = UAYV^* + UXBV^* <=> T = UAU^* [(U^*)^{-1} YV^*] + UXV^{-1} [VBV^*] <=> \tilde{T} = S(A)\tilde{Y} + \tilde{X} S(B)$, where $\tilde{X} = UXV^{-1}$, $\tilde{Y} = (U^*)^{-1} YV^*$. It follows that

$$\varphi \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} <=> \varphi \begin{bmatrix} S(A) & \tilde{T} \\ 0 & S(B) \end{bmatrix}.$$

Hence setting $T = \tilde{T}$, we may assume w.l.o.g. that $A = S(A)$, $B = S(B)$.

(ii) Let $r' = \text{rank } A$, $s' = \text{rank } B$. We may assume w.l.o.g. that $T = (t_{i,r+j})_{\substack{1 \le i \le r \\ 1 \le j \le s}}$, where $t_{i,r+j} = 0$ for (i,j) such that $r' < i \le r$ or $s' < i \le s$. Justification: Let $A = S(A) = \text{diag}(\alpha_1, \ldots, \alpha_{r'}, 0, \ldots, 0)$, $B = S(B) = \text{diag}(\beta_1, \ldots, \beta_{s'}, 0, \ldots, 0)$, where $\alpha_1 | \alpha_2 | \ldots | \alpha_{r'}$ and $\beta_1 | \beta_2 | \ldots | \beta_{s'}$. Assume first that $r' < r$ and $s' < s$. If $\exists\, (i,j) \in (r', r] \times (s', s]$ such that $t_{i,r+j} \ne 0$, then it would follow that $\text{rank} \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} > r' + s' = \text{rank} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, a contradiction. Hence $t_{i,r+j} = 0$ for $(i,j) \in (r', r] \times (s', s]$.

Assume now that $r' < r$ and choose $(i, j) \in (r', r] \times [1, s']$. Then it is easily seen that $d_{r'+s'} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = \alpha_1 \ldots \alpha_{r'} \beta_1 \ldots \beta_{s'}$, and that

$$\delta = \begin{cases} \alpha_1 \ldots \alpha_{r'} \beta_1 \ldots \beta_{j-1} \beta_{j+1} \ldots \beta_{s'} t_{i,r+j}, & j < s' \\ \alpha_1 \ldots a_{r'} \beta_1 \ldots \beta_{j-1} t_{i,r+j} & , j = s' \end{cases}$$

is an $(r'+s') \times (r'+s')$ determinantal divisor of $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix}$. Since $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$, it follows that $\alpha_1 \ldots \alpha_{r'} \beta_1 \ldots \beta_{s'} | \delta$, from which we deduce that $\beta_j | t_{i,r+j}$. Hence we may choose $w_{i,r+j} \in R$ such that $t_{i,r+j} = w_{i,r+j} \beta_j$.

Assume finally that $s' < s$ and choose $(i,j) \in [1, r'] \times (s', s]$. Then it is easily seen that

$$\eta = \begin{cases} \alpha_1 \ldots \alpha_{i-1} \alpha_{i+1} \ldots \alpha_{r'} \beta_1 \ldots \beta_{s'} t_{i,r+j}, & i < r' \\ \alpha_1 \ldots \alpha_{i-1} \beta_1 \ldots \beta_{s'} t_{i,r+j} & , i = r' \end{cases}$$

is an $(r'+s') \times (r'+s')$ determinantal divisor of $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix}$. It follows that $\alpha_1 \ldots \alpha_{r'} \beta_1 \ldots \beta_{s'} | \eta$, from which we deduce that $\alpha_i | t_{i,r+j}$. Hence we may choose $z_{i,r+j} \in R$ such that $t_{i,r+j} = \alpha_i z_{i,r+j}$.

Now for $1 \le i \le r$, $1 \le j \le s$ set

$$\bar{w}_{i,r+j} = \begin{cases} w_{i,r+j}, \text{ if } r' < i \le r,\ 1 \le j \le s',\ t_{i,r+j} \ne 0 \\ 0 \quad \text{otherwise} \end{cases}$$

and set

$$\bar{z}_{i,r+j} = \begin{cases} z_{i,r+j} \text{ if } 1 \le i \le r',\ s' < j \le s,\ t_{i,r+j} \ne 0 \\ 0 \quad \text{otherwise.} \end{cases}$$

Let $W = (\bar{w}_{i,r+j})_{\substack{1 \le i \le r \\ 1 \le j \le s}}$, $Z = (\bar{z}_{i,r+j})_{\substack{1 \le i \le r \\ 1 \le j \le s}} \in R_{r,s}$.

Then $\begin{bmatrix} I_r & -W \\ 0 & I_s \end{bmatrix} \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \begin{bmatrix} I_r & -Z \\ 0 & I_s \end{bmatrix} = \begin{bmatrix} A & \tilde{T} \\ 0 & B \end{bmatrix}$, where $\tilde{T} = T - AZ - WB$. Note that

$$(\tilde{T})_{i,r+j} = \begin{cases} T_{i,r+j}, & 1 \leqslant i \leqslant r' \text{ or } 1 \leqslant j \leqslant s' \\ 0 & \text{otherwise.} \end{cases}$$

It follows that $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} A & \tilde{T} \\ 0 & B \end{bmatrix}$, so that $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} <=> \begin{bmatrix} A & \tilde{T} \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$. It is also immediate that $T = AY + XB \ \langle = \rangle\ \tilde{T} = A\tilde{Y} + \tilde{X}B$, where $\tilde{X} = X - W$, $\tilde{Y} = Y - Z$. Thus $\varphi \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \langle = \rangle\ \varphi \cdot \begin{bmatrix} A & \tilde{T} \\ 0 & B \end{bmatrix}$. Here we may assume w.l.o.g. that $T = \tilde{T}$, i.e., that T is of the form specified above.

(iii) We may assume w.l.o.g. that rank $A = r$, rank $B = s$. Justification: We have from (i) and (ii) that we may assume that $A = \tilde{A} \dotplus 0_{r-r'}$, where $\tilde{A} = \mathrm{diag}\ (\alpha_1, \ldots, \alpha_{r'})$, $B = \tilde{B} \dotplus 0_{s-s'}$, where $\tilde{B} = \mathrm{diag}\ (\beta_1, \ldots, \beta_{s'})$, and that $T = \tilde{T} \dotplus 0_{r-r',\, s-s'}$, where $\tilde{T} \in R_{r',s'}$. It is not difficult to show that

$$\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \langle = \rangle \begin{bmatrix} \tilde{A} & \tilde{T} \\ 0 & \tilde{B} \end{bmatrix} \tilde{E} \begin{bmatrix} \tilde{A} & 0 \\ 0 & \tilde{B} \end{bmatrix},$$

also that $T = AY + XB$ for some $X, Y \in R_{rs}$ $\langle = \rangle$ $\tilde{T} = \tilde{A}\tilde{Y} + \tilde{X}\tilde{B}$ for some $\tilde{X}, \tilde{Y} \in R_{r's'}$. Thus

$$\varphi \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \langle = \rangle\ \varphi \begin{bmatrix} \tilde{A} & \tilde{T} \\ 0 & \tilde{B} \end{bmatrix}.$$

Hence we may assume w.l.o.g. that $A = \tilde{A}$, $B = \tilde{B}$, i.e., that rank $A = r$, *rank* $B = s$.

(iv) We may assume w.l.o.g. that $r = s$. Justification: Assume $r < s$. Let

$$\tilde{A} = I_{s-r} \dotplus A,\ T = \begin{bmatrix} 0_{s-r,s} \\ T \end{bmatrix}.$$

It is an easy consequence of [1, Ch. 2, ex. 1] that

$$\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \langle = \rangle \begin{bmatrix} \tilde{A} & \tilde{T} \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} \tilde{A} & 0 \\ 0 & B \end{bmatrix}.$$

It is also not difficult to show that $T = AY + XB$ for some $X, Y \in R_{rs}$ $\langle = \rangle$ $\tilde{T} = \tilde{A}\tilde{Y} + \tilde{X}B$ for some $\tilde{X}$, $\tilde{Y} \in R_{ss}$. Thus

$$\varphi \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \langle = \rangle\ \varphi \begin{bmatrix} \tilde{A} & \tilde{T} \\ 0 & B \end{bmatrix}.$$

Assume now that $s < r$. Let $\tilde{B} = I_{r-s} \dotplus B$, $\tilde{T} = [0_{r,r-s}, T]$. Proceeding as above, we can show that

$$\varphi \begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \langle = \rangle\ \varphi \begin{bmatrix} A & \tilde{T} \\ 0 & \tilde{B} \end{bmatrix}.$$

It then follows from this and the above case that we may assume w.l.o.g. that $r = s$.

We now complete the proof of the theorem. By (i)—(iv), we may assume w.l.o.g. that $A = S(A) = \mathrm{diag}\ (\alpha_1, \ldots, \alpha_r)$, $B = S(B) = \mathrm{diag}\ (\beta_1, \ldots, \beta_r)$, where α_i, $\beta_j \neq 0$, $1 \leqslant i$, $j \leqslant r$. Note that $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix} \tilde{\mathrm{E}} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \langle = \rangle\ \forall\, k \leqslant 2r$, $\mathrm{d}_k \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \Big|$ every $k \times k$ determinantal divisor of $\begin{bmatrix} A & T \\ 0 & \mathrm{B} \end{bmatrix}$. Note also that $T=AY+XB \langle=\rangle (\alpha_i, \beta_j) \mid t_{i,r+j}, 1 \leq i, j \leq r$. Now

$$d_k \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = (\alpha_1 \ldots \alpha_k, \beta_1 \ldots \beta_k, \{\alpha_{r-t}\alpha_{r-t+1} \ldots \alpha_r\beta_1 \ldots \beta_{k-(t+1)}\}_{1 \leq t \leq k-2}),$$

as is easily calculated. We consider two cases.

a. $i \geq j$. Let $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i}$ denote the matrix obtained by deleting row $r+j$ and column i from $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$. It is easily seen that if ρ is any $(i-1) \times (i-1)$ determinantal minor of $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i}$, then $\rho t_{i,r+j}$ is an $i \times i$ determinantal minor of $\begin{bmatrix} A & T \\ 0 & B \end{bmatrix}$. It then follows that

$$d_i \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} d_{i-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} t_{i,r+j}.$$

Note that

$$d_{i-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} = (\alpha_1 \ldots \alpha_{i-1}, \begin{cases} \beta_1 \ldots \beta_{j-1}\beta_{j+1} \ldots \beta_i, \text{if } j < i \\ \beta_1 \ldots \beta_{i-1}, \text{if } j = i \end{cases},$$

$$\{\alpha_1 \ldots \alpha_{i-u}\, \beta_1 \ldots \beta_{u-1}\}\, 2 \leq u \leq j,$$
$$\{\alpha_1 \ldots \alpha_{i-v}\, \beta_1 \ldots \beta_{j-1}\, \beta_{j+1} \ldots \beta_v\}_{j+1 \leq v \leq i-1}) \text{ (for } j \leq i-2).$$

From this it may be verified that

$$(\alpha_i, \beta_j)\, d_{i-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} \Bigg| d_i \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}.$$

It follows that $(a_i, \beta_j) \mid t_{i,r+j}$, as was to be shown.

b. $j > i$. We proceed as in (a). If σ is any $(j-1) \times (j-1)$ determinantal minor of $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i}$ then $\sigma t_{i,r+j}$ is a $j \times j$ determinantal minor of $\begin{bmatrix} A & T \\ 0 & \mathrm{B} \end{bmatrix}$. It then follows that

$$d_j \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \Bigg| d_{j-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} t_{i,r+j.}$$

Note that

$$d_{j-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} = (\alpha_1 \ldots \alpha_{i-1}\alpha_{i+1} \ldots \alpha_j,$$

$$\beta_1 \ldots \beta_{j-1}, \{\alpha_1 \ldots \alpha_u\beta_1 \ldots \beta_{j-(u+1)}\}_{1 \leq u \leq i-1},$$

$$\{\alpha_1 \ldots \alpha_{i-1}\alpha_{i+1} \ldots \alpha_v\beta_1 \ldots \beta_{j-v}\}_{i+1 \leq v \leq j-1}).$$
$$\text{(assuming } i \leq j+2)$$

From this it may be verified that $(\alpha_i,\beta_j)d_{j-1} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}_{r+j,i} \Bigg| d_j \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$.

It follows that $(\alpha_i,\beta_j) \mid t_{i,r+j}$, as was to be shown.
With this we have completed proof of the theorem. Q.E.D.

COROLLARY 1.1: *Let* A, B, *and* R *be as before, and suppose* P, Q$\in$R$_{rs}$.

$$\text{Then} \begin{bmatrix} A & Q-P \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \Rightarrow \begin{bmatrix} A & P \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} A & Q \\ 0 & B \end{bmatrix}.$$

PROOF: By Theorem 1, $Q-P=AY+XB$, for some $X, Y \in R_{rs}$. *Note that*

$$\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} A & P \\ 0 & B \end{bmatrix} \begin{bmatrix} I & Y \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & Q \\ 0 & B \end{bmatrix}. \text{ Hence } \begin{bmatrix} A & P \\ 0 & B \end{bmatrix} \tilde{E} \begin{bmatrix} A & Q \\ 0 & B \end{bmatrix} \qquad \text{Q.E.D.}$$

NOTE: The converse to Corollary 1.2 fails. For example, $\begin{bmatrix} 6 & 4 \\ 0 & 9 \end{bmatrix} \tilde{E} \begin{bmatrix} 6 & 2 \\ 0 & 9 \end{bmatrix}$, as may be verified by considering determinantal divisors, but $\begin{bmatrix} 6 & 4-2 \\ 0 & 9 \end{bmatrix}$ is *not* equivalent to $\begin{bmatrix} 6 & 0 \\ 0 & 9 \end{bmatrix}$.

We now generalize Theorem 1 as follows:

THEOREM 2: *Let* M *be a matrix over* R, *and suppose that* M *may be partitioned as*

$$M = \begin{bmatrix} M_{11} & M_{12} \dots & M_{1t} \\ 0 & M_{22} \dots & M_{2t} \\ \vdots & & \vdots \\ 0 \dots & 0 & M_{tt} \end{bmatrix}$$

where each M_{ii} *is square, of order* r_i. *Then* $M \tilde{E}$ *diag*$[M_{11}, \dots, M_{tt}]$ $\langle = \rangle$ *for* $1 \leq i < j \leq t$ $\exists$ X_{ij}, $Y_{ij} \in R_{r_i r_j}$ *such that* $M_{ij} = M_{ii} Y_{ij} + \sum_{k=i+1}^{j} X_{ik} M_{kj}$.

PROOF ($\Leftarrow$): Let

$$U = \begin{bmatrix} 1 & X_{12} & \dots & X_{1t} \\ 0 & 1 & 0 \dots & 0 \\ \vdots & & & \vdots \\ \cdot & & & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & & \dots 0 \\ 0 & 1 & X_{23} \dots & X_{2t} \\ \cdot & \cdot & & \cdot \\ \cdot & & & \cdot \\ 0 & \dots & 0 & 1 \end{bmatrix} \dots \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 \dots & 0 \dots & 0 \\ & & \vdots & \vdots \\ & & 1 & X_{t-1,t} \\ 0 \dots & & 0 & 1 \end{bmatrix}$$

and

$$V = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & 0 \dots & 0 \\ \vdots & & 1 & \vdots \\ & & & Y_{t-1,t} \\ 0 & \dots & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \dots & 0 & 0 \\ 0 & 1 & \vdots & \vdots \\ \vdots & & Y_{t-2,t-1} & Y_{t-2,t} \\ & & 1 & 0 \\ 0 \dots & 0 & & 1 \end{bmatrix} \begin{bmatrix} 1 & Y_{12} & \dots & Y_{1t} \\ 0 & 1 & 0 & 0 \\ \vdots & & & \vdots \\ \cdot & & & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}$$

Then U, V are invertible and U diag$[M_{11}, \dots, M_{tt}]$ $V = M$, as may be verified. Hence $M \tilde{E}$ diag $[M_{11}, \dots, M_{tt}]$.

($\Rightarrow$) Let $A_1 = M_{11}$, $T_1 = [M_{12}, \ldots, M_{1t}]$ and

$$B_1 = \begin{bmatrix} M_{22} & \ldots\ldots\ldots & M_{2t} \\ 0 & & \cdot \\ \vdots & & \vdots \\ 0 & \ldots\ldots\ldots & M_{tt} \end{bmatrix}$$

so that $M = \begin{bmatrix} A_1 & T_1 \\ 0 & B_1 \end{bmatrix}$. Note that to obtain a minor of $\begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix}$ with nonzero determinant, it is necessary that the number of rows deleted which pass through the block A_1 equal the number of columns deleted which pass through this block. It follows from this that every determinantal divisor of $\begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix}$ is also a determinantal divisor of M. Since $M \tilde{E}$ diag$[M_{11}, \ldots, M_{tt}]$, it follows that $\begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix} \tilde{E}$ diag$[M_{11}, \ldots, M_{tt}]$ as well, so that $M \tilde{E} \begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix}$. Hence by Theorem 1 there are matrices X_1, Y_1 of appropriate size such that $T_1 = A_1 Y_1 + X_1 B$. We may write $X_1 \equiv [X_{12}, \ldots, X_{1t}]$, $Y_1 = [Y_{12}, \ldots, Y_{1t}]$, where $X_{1j}, Y_{1j} \epsilon R_{r_1, rj}$, $2 \leq j \leq t$, from which it follows that $M_{1j} = M_{11} Y_{1j} + \sum_{k=2}^{j} X_{1k} M_{kj}$.

Now let $A_2 = M_{22}$, $T_2 = [M_{23}, \ldots, M_{2t}]$,

$$B_2 = \begin{bmatrix} M_{33} & \ldots & & M_{3t} \\ 0 & & & \\ \vdots & & & \vdots \\ 0 & \ldots & 0 & M_{tt} \end{bmatrix}$$

Then $M \tilde{E} A_1 \dot{+} \begin{bmatrix} A_2 & T_2 \\ 0 & B_2 \end{bmatrix} \tilde{E} A_1 \dot{+} \begin{bmatrix} A_2 & 0 \\ 0 & B_2 \end{bmatrix}$ as before, so that $\begin{bmatrix} A_2 & T_2 \\ 0 & B_2 \end{bmatrix} \tilde{E} \begin{bmatrix} A_2 & 0 \\ 0 & B_2 \end{bmatrix}$. Proceeding as above we obtain that $\exists\, X_2 \equiv [X_{23}, \ldots, X_{2t}]$, $Y_2 = [Y_{23}, \ldots, Y_{2t}]$ such that

$$M_{2j} = M_{22} Y_{2j} + \sum_{k=3}^{j} X_{2k} M_{kj},\ 3 \leq j \leq t.$$

Continuing in this manner we obtain the desired linear recurrence relation.

Q.E.D.

COROLLARY 2.1: *Let* R *and* $\{M_{ii}\}_{i=1}^{t}$ *be as in Theorem 2, and*

$$M \equiv \begin{bmatrix} M_{11} & M_{12} & \ldots & & M_{1t} \\ 0 & M_{22} & \ldots & & M_{2t} \\ \vdots & & & & \vdots \\ 0 & \ldots\ldots\ldots & & 0 & M_{tt} \end{bmatrix}$$

Then $M \tilde{E}$ *diag*$[M_{11}, M_{22}, \ldots, M_{tt}]\ \forall\ \{M_{ij}\}_{1 \leq i < j \leq t}$, *where* $M_{ij} \epsilon R_{r_i r_j}$ $\langle = \rangle$

(i) $(det\, M_{ii}, det\, M_{jj}) = 1$, $1 \leq i \leq j \leq t$ *or*

(ii) $\exists j \epsilon [1, n]$ *such that det* $M_{jj} = 0$ *and such that for all* $k \neq j$, M_{kk} *is unimodular.*

PROOF: ($\Leftarrow$) (i) This is essentially Theorem 3 of [1]. (ii) Consider first the equation

$$M_{t-1,t}=M_{t-1,t-1}Y_{t-1,t}+X_{t-1,t}M_{tt}. \quad (*)$$

By hypothesis, at least one of $M_{t-1,t-1}$, M_{tt} is unimodular. If $M_{t-1,t-1}$ is unimodular, we may let $X_{t-1,t}=0$ and $Y_{t-1,t}=M_{t-1,t-1}{}^{-1}\, M_{t-1,t}$ to obtain a solution to (*). If M_{tt} is unimodular, it is again easy to solve (*).

Consider now the equations

$$M_{t-2,t-1}=M_{t-2,t-2}Y_{t-2,t-1}+X_{t-2,t-1}M_{t-1,t-1} \quad (**)$$

and

$$M_{t-2,t}=M_{t-2,t-2}Y_{t-2,t}+X_{t-2,t-1}M_{t-1,t}+X_{t-2,t}M_{tt} \quad (***)$$

Proceeding as above, it is again easy to solve (**), this time for $X_{t-2,t-1}$, $Y_{t-2,t-1}$. Now rewrite (***) as

$$M_{t-2,t}-X_{t-2,t-1}M_{t-1,t}=M_{t-2,t-2}Y_{t-2,t}+X_{t-2,t}M_{tt} \quad (***')$$

and note that the matrices on the left-hand side have all been given or determined previously. Again, it is easy to solve (***'), for $X_{t-2,t}$, $Y_{t-2,t}$.

Proceeding in this manner, for $1\leqslant i<j\leqslant t$ we may find X_{ij}, $Y_{ij}\epsilon R_{r_i r_j}$ such that $M_{ij}=M_{ij}Y_{ij}+\sum_{k=i+1}^{j} X_{ik}M_{kj}$. Hence by Theorem 2, $M\widetilde{E}$ diag$[M_{11}, M_{22}, \ldots, M_{tt}]$.

($\Rightarrow$) We may assume w.o.l.g. that $M_{ii}=S(M_{ii})=\text{diag}(\alpha_{i1}, \ldots, \alpha_{ir_{i'}}, 0, \ldots, 0)$, $1\leqslant i\leqslant t$, where $r_{i'}\leqslant r_i$. Suppose $\exists$ distinct i, $j\epsilon[1, n]$ such that det $M_{ii}=$det $M_{jj}=0$. Then $r_i'<r_i$ and $r'<r_j$. But if we let M_{ij} be the matrix of all 1's and set $M_{uv}=0$, $u\neq i$ or $v\neq j$, we would obtain that rank $M>$ rank diag$[M_{11}, \ldots, M_{tt}]$, contradicting hypothesis. Hence there is at most one $i\epsilon[1,n]$ such that det $M_{ii}=0$.

Suppose first that there is such an i. We will show that (i) holds in this case. Choose any $j>i$ (if such exist) and let M_{ij} and M_{uv} be defined as above. By Theorem 2,

$$M_{ij}=M_{ii}Y_{ij}+\sum_{k=i+1}^{j} X_{ik}M_{kj}=M_{ii}Y_{ij}+X_{ij}M_{jj}.$$

Considering the (k, l) component of this matric equation, for $r_{i'}<k\leqslant r_i$ and $l\leqslant r_j$ we obtain that $1=(x_{ij})_{kl}\alpha_{jl}$. This implies that α_{jl} is a unit, $1\leqslant l\leqslant r_j$. Since det $M_{jj}=\prod_{i=1}^{rj}\alpha_{jl}$, we obtain that M_{jj} is unimodular. If we choose any $j<i$, we may obtain by a similar argument that det M_{jj} is a unit, so that M_{jj} is unimodular in this case as well. We can thus conclude that (i) holds when det $M_{ii}=0$ for some i.

Now suppose that det $M_{ii}\neq 0$, $1\leqslant i\leqslant t$. We will show that (ii) holds in this case. Choose $i\neq j\epsilon[1, t]$ and let M_{ij} and $M_{uv}(u\neq i$ or $v\neq j)$ be defined as before. Again, $M_{ij}=M_{ii}Y_{ij}+X_{ij}M_{jj}$. Considering the (k, l) component of this matric equation, for $1\leqslant k\leqslant r_i$, $1\leqslant l\leqslant r_j$ we obtain that

$$1=\alpha_{ij}(Y_{ij})_{kl}+(X_{ij})_{kl}\alpha_{jl}.$$

It follows that $(\alpha_{ik}, \alpha_{jl}) = 1$. Since $\det M_{ii} = \prod_{k=1}^{r_i} \alpha_{ik}$ and $\det M_{jj} = \prod_{l=1}^{r_j} \alpha_{jl}$, we obtain that $(\det M_{ii}, \det M_{ij}) = 1$. We can thus conclude that (ii) holds when $\det M_{ii} \neq 0$, all i.

This completes the proof of the corollary.

Q.E.D.

References

[1] Newman, Morris, Integral Matrices (Academic Press, New York, 1972).

[2] Newman, Morris, The Smith normal form of a partitioned matrix, J. Res. Nat. Bur. of Stand. (U.S.), **75B** (Math. Sci.), No. 1, 3–6 (Jan.–March 1974).

[3] Roth, W., The equations $AX - YB = C$ and $AX - XB = C$ in matrices, Proc. AMS **3**, 392–396 (1952).

(Paper 80B1–435)

211-694 O - 76 - 7