On the Matrices AB and BA*

R. C. THOMPSON
University of California
Santa Barbara, California

Communicated by Marvin Marcus

1. INTRODUCTION

Throughout this paper A, B, P, Q are, respectively, $m \times n$, $n \times m$, $m \times m$, $n \times n$ matrices with elements in some given field K. Let $P = AB$ and $Q = BA$. In a well-known paper [1], H. Flanders solved the problem of determining the relationship between the elementary divisors of P and those of Q. Another proof of Flanders' theorem, with some generalizations, has been given in [2]. In this paper we give a new proof of Flanders' theorem. It is obvious that some connection exists between the ranks of A and B and the intertwining of the elementary divisors of P and Q. Our proof enables us to derive a set of inequalities relating the integers rank A, rank B to the elementary divisor structures of P and Q. These inequalities are stronger than the obvious inequalities rank P, rank $Q \leqslant$ rank A, rank $B \leqslant m, n$. Our main result is that our inequalities are the necessary and sufficient conditions in order that $P = AB$ and $Q = BA$ where A and B are to have prescribed ranks. Our inequalities define an isosceles trapezoid of permissible values for the integer pairs (rank A, rank B) for fixed P, Q. We investigate the relationship between properties of P, Q and properties of the associated trapezoid. We go on to determine which isosceles trapezoids are achievable as trapezoids of permissible values for some pair P, Q. Finally we investigate when a given trapezoid can be associated with a pair P, Q for which $P = Q$. Additional results are also derived.

* The preparation of this paper was supported in part by the U.S. Air Force under Grant 698-67.

Linear Algebra and Its Applications 1, 43—58 (1968)

Our methods use only elementary matrix theory, and are thus accessible to any linear algebra student familiar with elementary divisor theory and the Hermite (echelon) method for solving linear systems.

We define the degree function by degree $P = m$ for P an $m \times m$ matrix.

2. NECESSITY

From

$$P = AB, \qquad Q = BA, \tag{1}$$

we get

$$SPS^{-1} = (SAT)(T^{-1}BS^{-1}), \qquad T^{-1}QT = (T^{-1}BS^{-1})(SAT), \tag{2}$$

where S is $m \times m$ and T is $n \times n$. Thus in (1) we may make independent similarities of P and Q or, alternatively, an equivalence of A and B. Thus we may assume $A = I_r \dotplus 0_{m-r,n-r}$, where I_r denotes the r-square identity matrix, $0_{\alpha\beta}$ is an $\alpha \times \beta$ matrix of zeros, and $\dotplus$ denotes direct sum. We partition

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \tag{3}$$

conformally with the partitioning of A. We proceed to simplify B, step by step, retaining the form of A.

Let $S = S_1 \dotplus I_{m-r}$ and $T = S_1^{-1} \dotplus I_{n-r}$. Then, as in (2), B_{11} is converted to $S_1 B_{11} S_1^{-1}$. Hence in (3) we may assume $B_{11} = \bar{B}_{11} \dotplus \tilde{B}_{11}$, where $\bar{B}_{11}$ is a nonsingular $f \times f$ matrix, and $\tilde{B}_{11}$ is an $e \times e$ nilpotent matrix in Jordan form: $\tilde{B}_{11} = J_{e_1} \dotplus J_{e_2} \dotplus \cdots \dotplus J_{e_t}$, where $e = e_1 + e_2 + \cdots + e_t$, and J_α denotes the α-square Jordan matrix belonging to elementary divisor $\lambda^\alpha : J_\alpha$ is all zeros except for a stripe of ones on the diagonal just above the main diagonal. We arrange the blocks in $\tilde{B}_{11}$ so that $e_1 \geqslant e_2 \geqslant \cdots \geqslant e_t$.

Now let

$$B_{12} = \begin{bmatrix} \bar{B}_{12} \\ \tilde{B}_{12} \end{bmatrix}, \qquad B_{21} = [\bar{B}_{21}, \tilde{B}_{21}].$$

Here $\bar{B}_{12}$ is $f \times (m-r)$, $\tilde{B}_{12}$ is $e \times (m-r)$, $\bar{B}_{21}$ is $(n-r) \times f$, $\tilde{B}_{21}$ is $(n-r) \times e$. Let

$$S = \begin{bmatrix} I_f & 0 & \bar{B}_{11}^{-1}\bar{B}_{12} \\ 0 & I_e & 0 \\ 0 & 0 & I_{m-r} \end{bmatrix}, \qquad T = \begin{bmatrix} I_f & 0 & 0 \\ 0 & I_e & 0 \\ \bar{B}_{21}\bar{B}_{11}^{-1} & 0 & I_{n-r} \end{bmatrix}.$$

Changing notation and denoting SAT by A and $T^{-1}BS^{-1}$ by B, we now find that $\bar{B}_{12} = 0$ and $\bar{B}_{21} = 0$.

Next let

$$S = I_f \dotplus \begin{bmatrix} I_e & W \\ 0 & I_{m-r} \end{bmatrix}, \qquad T = I_f \dotplus \begin{bmatrix} I_e & 0 \\ Z & I_{n-r} \end{bmatrix}.$$

Then for $T^{-1}BS^{-1}$ we still have $\bar{B}_{12} = 0$, $\bar{B}_{21} = 0$, and now block $\tilde{B}_{12}$ is replaced with $-\tilde{B}_{11}W + \tilde{B}_{12}$ and block $\tilde{B}_{21}$ is replaced with $-Z\tilde{B}_{11} + \tilde{B}_{12}$. Thus in $\tilde{B}_{12}$ we can, choosing W appropriately, add an arbitrary linear combination of the columns of $\tilde{B}_{11}$ to each given column of $\tilde{B}_{12}$, and, choosing Z appropriately, add an arbitrary linear combination of the rows of $\tilde{B}_{11}$ to each given row of $\tilde{B}_{21}$. This means that in B we may suppose that each row passing through a one in a Jordan block of $\tilde{B}_{11}$ consists of zeros only when intersecting $\tilde{B}_{12}$, and that each column of B passing through a one in a Jordan block of $\tilde{B}_{11}$ consists of zeros only when intersecting $\tilde{B}_{21}$.

Continuing our simplification of B, we next let $S = I_f \dotplus I_e \dotplus W_1$ and $T = I_f \dotplus I_e \dotplus Z_1$, where W_1 and Z_1 are nonsingular. Then in $T^{-1}BS^{-1}$ block $\tilde{B}_{12}$ becomes $\tilde{B}_{12}W_1^{-1}$ and block $\tilde{B}_{21}$ becomes $Z_1^{-1}\tilde{B}_{21}$. Thus we may assume that $\tilde{B}_{12}$ is in column Hermite (echelon) form and $\tilde{B}_{21}$ is in row Hermite form. The zero rows of $\tilde{B}_{12}$ and the zero columns of $\tilde{B}_{21}$ obtained in the previous paragraph are preserved.

Thus let $\tilde{B}_{12} = (b_{ij})_{1 \leqslant i \leqslant t, 1 \leqslant j \leqslant m-r}$, where b_{ij} is a column e_i-tuple with all positions zero except perhaps for the bottom position. Because $\tilde{B}_{12}$ is in Hermite form we may suppose that the nonzero columns of $\tilde{B}_{12}$ are columns $1, 2, \ldots, s$, that $b_{1i} = 0, b_{2i} = 0, \ldots, b_{c_i-1,i} = 0, b_{c_i i} \neq 0$, that $c_1 < c_2 < \cdots < c_s$, that the nonzero entry of $b_{c_i i}$ is a one, and that $b_{c_i 1} = b_{c_i 2} = \cdots = b_{c_i, i-1} = 0$; $1 \leqslant i \leqslant s$. Similarly let $\tilde{B}_{21} = (\beta_{ij})_{1 \leqslant i \leqslant n-r, 1 \leqslant j \leqslant t}$, where β_{ij} is a row e_j-tuple, with all entries zero except perhaps for the first entry. Because $\tilde{B}_{21}$ is in Hermite form, we may assume that the nonzero rows of $\tilde{B}_{21}$ are rows $1, 2, \ldots, u$, that $\beta_{i1} = 0$, $\beta_{i2} = 0, \ldots, \beta_{i,d_i-1} = 0$, $\beta_{id_i} \neq 0$, that $d_1 < d_2 < \cdots < d_u$, that the nonzero entry of β_{id_i} is a one, and that $\beta_{1d_i} = \beta_{2d_i} = \cdots = \beta_{i-1,d_i} = 0$; $1 \leqslant i \leqslant u$.

Summarizing, we have $A = I_f \dotplus I_e \dotplus 0_{m-e-f,n-e-f}$ and

$$B = \begin{bmatrix} \bar{B}_{11} & 0 & 0 \\ 0 & \tilde{B}_{11} & \tilde{B}_{12} \\ 0 & \tilde{B}_{21} & B_{22} \end{bmatrix}, \tag{4}$$

where $\tilde{B}_{11}$, $\tilde{B}_{12}$, $\tilde{B}_{21}$ satisfy the conditions obtained above. Then we get

$$P = \begin{bmatrix} \bar{B}_{11} & 0 & 0 \\ 0 & \tilde{B}_{11} & \tilde{B}_{12} \\ 0 & 0 & 0 \end{bmatrix}, \qquad Q = \begin{bmatrix} \bar{B}_{11} & 0 & 0 \\ 0 & \tilde{B}_{11} & 0 \\ 0 & \tilde{B}_{21} & 0 \end{bmatrix}. \tag{5}$$

Our next goal is to compute the elementary divisors of P and of Q. To this end we shall carry out certain similarity transformations of P and, later, of Q.

Make the similarity transformation of P (a sequence of column exchanges and the same sequence of row exchanges) that moves the column of P containing column i of $\tilde{B}_{12}$ to a new column position between $J_{e_{c_i}}$ and $J_{e_{c_i+1}}$; the corresponding row operation moves a zero row up to a new row position between $J_{e_{c_i}}$ and $J_{e_{c_i+1}}$; $1 \leqslant i \leqslant s$. The new matrix, call it X, has the form of a block triangular matrix in which the main block diagonal is

$$\bar{B}_{11} \dotplus J_{e_1} \dotplus \cdots \dotplus J_{e_{c_1}+1} \dotplus \cdots \dotplus J_{e_{c_s}+1} \dotplus \cdots \dotplus J_{e_t} \dotplus 0_{m-r-s,m-r-s}.$$

An off-diagonal block in X can be nonzero only if it appears in the same block column as some $J_{e_{c_i}+1}$ and in the same block row as some J_{e_ρ} with $c_i < \rho \leqslant t$, $\rho \neq c_{i+1}, \ldots, \rho \neq c_s$. Call this nonzero block $X_{\rho i}$. Note that the block row of $J_{e_{c_i}+1}$ is entirely zero except for the main block diagonal position as is the block column of J_{e_ρ}. Let $x_{\rho i}$ be the nonzero element in $X_{\rho i}$; $x_{\rho i}$ is in the lower right corner of $X_{\rho i}$. Since $\rho > i$, $e_{c_i} + 1 > e_\rho$. Let $W_{\rho i}$ be $e_\rho \times (e_{c_i} + 1)$ with all entries zero except for $-x_{\rho i}$ occupying all positions (μ, ν) for which $\mu - \nu = e_\rho - e_{c_i}$. Let S be block triangular with main block diagonal

$$I_f \dotplus I_{e_1} \dotplus \cdots \dotplus I_{e_{c_1}+1} \dotplus \cdots \dotplus I_{e_{c_s}+1} \dotplus \cdots \dotplus I_{e_t} \dotplus I_{m-r-s}.$$

All the off-diagonal blocks in S are to be zero, except for the blocks $W_{\rho i}$, where $W_{\rho i}$ sits in the block row of I_{e_ρ} and the block column of $I_{e_{c_i}+1}$. Then in SXS^{-1} all block positions are undisturbed, except that now in place of $X_{\rho i}$ we have $X_{\rho i} + W_{\rho i}J_{e_{c_i}+1} - J_{e_\rho}W_{\rho i} = 0$. Thus SXS^{-1} is in a form from which the elementary divisors can be read off. Hence the elementary divisors of P are the elementary divisors of $\bar{B}_{11}$ together with the elementary divisors

$$\lambda^{e_1}, \ldots, \lambda^{e_{c_1}+1}, \ldots, \lambda^{e_{c_s}+1}, \ldots, \lambda^{e_t}, \underbrace{\lambda, \ldots, \lambda}_{m-r-s}. \tag{6}$$

We now repeat this calculation with Q. Make the similarity transformation of Q that moves row i of $\tilde{B}_{21}$ to a new row position between $J_{e_{d_i}-1}$ and $J_{e_{d_i}}$; the corresponding column operation moves a column of zeros to a new column position between $J_{e_{d_i}-1}$ and $J_{e_{d_i}}$; $1 \leqslant i \leqslant u$. The new matrix, call it Y, is a block triangular matrix in which the main block diagonal is

$$\bar{B}_{11} \dotplus J_{e_1} \dotplus \cdots \dotplus J_{e_{d_1}+1} \dotplus \cdots \dotplus J_{e_{d_u}+1} \dotplus \cdots \dotplus J_{e_t} \dotplus 0_{n-r-u,\, n-r-u}.$$

An off-diagonal block in Y can be nonzero only if it appears in the same block row as some $J_{e_{d_i}+1}$ and in the same block column as some J_{e_ρ} with $\rho > d_i$, $\rho \leqslant t$, $\rho \neq d_{i+1}, \ldots, \rho \neq d_u$. Call this nonzero block $Y_{i\rho}$. Note that the block column of $J_{e_{d_i}+1}$ is entirely zero except for the main block diagonal position as is the block row of J_{e_ρ}. Let $y_{i\rho}$ be the nonzero element of $Y_{i\rho}$; $y_{i\rho}$ is in the upper left corner of $Y_{i\rho}$. Since $\rho > i$, $e_\rho < e_{d_i} + 1$. Let $V_{i\rho}$ be $(e_{d_i} + 1) \times e_\rho$ with all entries zero except for $-y_{i\rho}$ occupying all positions (μ, ν) in $V_{i\rho}$ for which $\mu - \nu = 1$. Let T be block triangular with main block diagonal

$$I_f \dotplus I_{e_1} \dotplus \cdots \dotplus I_{e_{d_1}+1} \dotplus \cdots \dotplus I_{e_{d_u}+1} \dotplus \cdots \dotplus I_{e_t} \dotplus I_{n-r-u}.$$

All off-diagonal blocks in T are to be zero except for the blocks $V_{i\rho}$ in the block row of $I_{e_{d_i}+1}$ and the block column of I_{e_ρ}. Then in $T^{-1}YT$ all block positions are undisturbed except that now $Y_{i\rho}$ is replaced with $Y_{i\rho} + J_{e_{d_i}+1}V_{i\rho} - V_{i\rho}J_{e_\rho} = 0$. Thus $T^{-1}YT$ is in a form from which the elementary divisors can be read off. Hence the elementary divisors of Q are the elementary divisors of $\bar{B}_{11}$ together with the elementary divisors

$$\lambda^{e_1}, \ldots, \lambda^{e_{d_1}+1}, \ldots, \lambda^{e_{d_u}+1}, \ldots, \lambda^{e_t}, \underbrace{\lambda, \ldots, \lambda}_{n-r-u}. \tag{7}$$

Rearrange the sequence (6) so that the exponents are nonincreasing; let this rearranged sequence be

$$\lambda^{m_1}, \lambda^{m_2}, \ldots, \lambda^{m_p}; \qquad m_1 \geqslant m_2 \geqslant \cdots \geqslant m_p > 0. \tag{8}$$

Rearrange the sequence (7) so that the exponents are nonincreasing; let this rearranged sequence be

$$\lambda^{m_1'}, \lambda^{m_2'}, \ldots, \lambda^{m_q'}; \qquad m_1' \geqslant m_2' \geqslant \cdots \geqslant m_q' > 0. \tag{9}$$

Because $e_1 \geqslant e_2 \geqslant \cdots \geqslant e_t$, it follows that

$$\begin{aligned} &|m_i - m_i'| \leqslant 1, \qquad 1 \leqslant i \leqslant \min(p, q); \\ &m_i = 1 \qquad \text{for} \quad i > p \quad (\text{if } p > q), \\ &m_i' = 1 \qquad \text{for} \quad i > q \quad (\text{if } q > p). \end{aligned} \tag{10}$$

For suppose that the distinct integers among $e_1, \ldots, e_t$ are $\varepsilon_1 > \cdots > \varepsilon_\varphi$ with multiplicities $\mu_1, \ldots, \mu_\varphi$, respectively. Suppose that the number of integers $c_1, \ldots, c_s$ within the interval $[\mu_1 + \cdots + \mu_{i-1} + 1, \mu_1 + \cdots + \mu_i]$ is δ_i and that the number of integers $d_1, \ldots, d_\mu$ within this interval is θ_i; $1 \leqslant i \leqslant \varphi$. Then the exponents in (6), when (6) is rearranged so that the exponents are nonincreasing, become

$$\begin{gathered} \varepsilon_1 + 1 \ (\delta_1 \text{ times}), \quad \varepsilon_1 \ (\mu_1 - \delta_1 \text{ times}), \quad \varepsilon_2 + 1 \ (\delta_2 \text{ times}), \\ \varepsilon_2 \ (\mu_2 - \delta_2 \text{ times}), \ldots, \quad \varepsilon_\varphi + 1 \ (\delta_\varphi \text{ times}), \quad \varepsilon_\varphi \ (\mu_\varphi - \delta_\varphi \text{ times}), \\ 1 \ (m - r - s \text{ times}); \end{gathered} \tag{11}$$

and the exponents in (7), when (7) is rearranged so that the exponents are nonincreasing, become

$$\begin{gathered} \varepsilon_1 + 1 \ (\theta_1 \text{ times}), \quad \varepsilon_1 \ (\mu_1 - \theta_1 \text{ times}), \quad \varepsilon_2 + 1 \ (\theta_2 \text{ times}), \\ \varepsilon_2 \ (\mu_2 - \theta_2 \text{ times}), \ldots, \quad \varepsilon_\varphi + 1 \ (\theta_\varphi \text{ times}), \quad \varepsilon_\varphi \ (\mu_\varphi - \theta_\varphi \text{ times}), \\ 1 \ (n - r - u \text{ times}). \end{gathered} \tag{12}$$

This completes the proof of half of Flanders' result.

THEOREM 1. *If $P = AB$ and $Q = BA$ then P and Q have the same elementary divisors belonging to nonzero eigenvalues, and the elementary divisors* (8) *of P and* (9) *of Q belonging to eigenvalue zero satisfy* (10).

Before continuing with our argument, we require some additional notation.

Notation: Referring to the elementary divisors (8) and (9) of P and Q respectively, and noting (10), we let:

$\#(P/Q)$ denote the number of integers i with $1 \leqslant i \leqslant \min(p, q)$ for which $m_i > m_i'$;

$\#(Q/P)$ denote the number of integers i with $1 \leqslant i \leqslant \min(p, q)$ for which $m_i' > m_i$;

$\eta(P, Q)$ denote the number of integers i with $1 \leqslant i \leqslant \min(p, q)$ for which $\max(m_i, m_i') > 1$;

$N(P, Q)$ denote the number of integers i with $1 \leqslant i \leqslant \min(p, q)$ for which $m_i = m_i' > 1$;

$L(P, Q)$ denote the number of integers i with $1 \leqslant i \leqslant \min(p, q)$ for which $m_i = m_i' = 1$.

Thus N denotes the number of nonlinear coincidences of elementary divisors for zero of P and Q, L denotes the number of linear coincidences, η denotes the total number of pairs excluding linear pairs, and $\#(Q/P) + \#(P/Q)$ the number of nonlinear disagreements. Also $\eta(P, Q) = N(P, Q) + \#(P/Q) + \#(Q/P)$.

We now continue the proof started above. Observe that in B given by (4) where $\bar{B}_{11}$, $\tilde{B}_{11}$, $\tilde{B}_{12}$, $\tilde{B}_{21}$ are as described above, the following sets of rows are an independent set: the rows of B passing through $\bar{B}_{11}$, together with the rows of B passing through the ones of the Jordan blocks in $\tilde{B}_{11}$, together with the rows of B passing through the leading ones in the nonzero columns of $\tilde{B}_{12}$, together with the rows of B passing through the leading ones in the nonzero rows of $\tilde{B}_{21}$. This independence stems from the fact that the ones in $\tilde{B}_{11}$, the leading ones in $\tilde{B}_{12}$, and the leading ones in $\tilde{B}_{21}$ all occur in different rows and in different columns. Thus rank $B \geqslant f + e - t + s + u$. Knowing the elementary divisors (6) and (7) of P and Q, we compute that rank $P = f + e - t + s$ and rank $Q = f + e - t + u$. We also have rank $A = f + e$. We also observe that the rows of B that pass through zero rows of $\tilde{B}_{11}$ and do not pass through leading ones of $\tilde{B}_{12}$ are dependent on the rows of B that pass both through zero rows of $\tilde{B}_{11}$ and through leading ones of $\tilde{B}_{12}$. Thus a maximal independent set of rows of B is obtained by adjoining some of the rows of B passing through $\tilde{B}_{21}$ to the following set of rows of B: the rows passing through $\bar{B}_{11}$, together with the rows passing through ones of Jordan blocks in $\tilde{B}_{11}$, together with rows passing through leading ones of $\tilde{B}_{12}$. Hence $\operatorname{rank} B \leqslant f + e - t + s + n - r = n - t + s$. An analogous column argument shows that the columns of B passing through the columns of $\bar{B}_{11}$, together with columns passing through ones in Jordan blocks of $\tilde{B}_{11}$, together with columns passing through leading ones in $\tilde{B}_{21}$, together with certain of the columns passing through $\tilde{B}_{12}$, form a maximal independent set of columns of B. Hence $\operatorname{rank} B \leqslant m - t + u$. We are now ready to prove Theorem 2.

THEOREM 2. *Let $P = AB$ and $Q = BA$. Then,*

$$\operatorname{rank} A \geqslant \operatorname{rank} P + \#(Q/P) = \operatorname{rank} Q + \#(P/Q), \tag{13}$$

$$\operatorname{rank} B \geqslant \operatorname{rank} P + \#(Q/P) = \operatorname{rank} Q + \#(P/Q), \tag{14}$$

$$\operatorname{rank} A + \operatorname{rank} B \geqslant \operatorname{rank} P + \operatorname{rank} Q + \eta(P, Q), \tag{15}$$

$$\operatorname{rank} A + \operatorname{rank} B \leqslant \min\{\operatorname{rank} P + n, \operatorname{rank} Q + m\}. \tag{16}$$

Proof. From (11) and (12) we obtain $\#(P/Q) = \sum_{i=1}^{\varphi} \max(\delta_i - \theta_i, 0) \leqslant \sum_{i=1}^{\varphi} \delta_i = s$. Also $\#(P/Q) = \sum_{i=1}^{\varphi} \max(\delta_i - \theta_i, 0) \leqslant \sum_{i=1}^{\varphi} \max(\mu_i - \theta_i, 0) = \sum_{i=1}^{\varphi} \mu_i - \sum_{i=1}^{\varphi} \theta_i = t - u$. Furthermore, $\#(Q/P) = \sum_{i=1}^{\varphi} \max(\theta_i - \delta_i, 0) \leqslant \sum_{i=1}^{\varphi} \theta_i = u$. Also $\#(Q/P) = \sum_{i=1}^{\varphi} \max(\theta_i - \delta_i, 0) \leqslant \sum_{i=1}^{\varphi} \max(\mu_i - \delta_i, 0) = t - s$. Lastly, $\eta \leqslant t$. Now note that $\operatorname{rank} B \geqslant f + e - t + u + s = \operatorname{rank} P + u \geqslant \operatorname{rank} P + \#(Q/P)$. This proves the inequality in (14). Next, observe $\operatorname{rank} B \geqslant f + e - t + u + s = \operatorname{rank} P + \operatorname{rank} Q - f - e + t = \operatorname{rank} P + \operatorname{rank} Q - \operatorname{rank} A + t \geqslant \operatorname{rank} P + \operatorname{rank} Q - \operatorname{rank} A + \eta$. From this, (15) follows. We also have $\operatorname{rank} A = f + e = f + e - t + s + t - s = \operatorname{rank} P + t - s \geqslant \operatorname{rank} P + \#(Q/P)$. This proves half of (13). Next, from $\operatorname{rank} B \leqslant n - t + s$ we get $\operatorname{rank} A + \operatorname{rank} B \leqslant f + e + n - t + s = \operatorname{rank} P + n$. Again $\operatorname{rank} A + \operatorname{rank} B \leqslant f + e + m - t + u = \operatorname{rank} Q + m$. This proves (16). The equality $\operatorname{rank} P + \#(Q/P) = \operatorname{rank} Q + \#(P/Q)$ follows from $\max(\delta_i - \theta_i, 0) + \theta_i = \max(\theta_i - \delta_i, 0) + \delta_i$. This completes the proof of Theorem 2.

3. SUFFICIENCY

We now let the elementary divisors of P and Q satisfy the conditions of Theorem 1. That is, the elementary divisors are a common set belonging to nonzero eigenvalues, with total degree f; for P, belonging to eigenvalue zero, a set (8) and for Q, belonging to eigenvalue zero, a set (9), such that (10) holds.

We change the notation used in Section 2, and now we define integer t by $\max(m_t, m_t') > 1$, $m_i = 1$ for all $i > t$ and $m_i' = 1$ for all $i > t$. Let $e_i = \min(m_i, m_i')$ for $1 \leqslant i \leqslant t$. Choose discrepancy numbers $c_1, \ldots, c_s$ and $d_1, \ldots, d_u$ so that the elementary divisors of P belonging to eigenvalue zero are

$$\lambda^{e_1}, \ldots, \lambda^{e_{c_1}+1}, \ldots, \lambda^{e_{c_s}+1}, \ldots, \lambda^{e_t}, \underbrace{\lambda, \ldots, \lambda}_{p-t}, \tag{17}$$

whereas those for Q belonging to eigenvalue zero are

$$\lambda^{e_1}, \ldots, \lambda^{e_{d_1}+1}, \ldots, \lambda^{e_{d_u}+1}, \ldots, \lambda^{e_t}, \underbrace{\lambda, \ldots, \lambda}_{q-t}. \tag{18}$$

Here we are arranging notation so that $c_1, \ldots, c_s$ are each distinct from each of $d_1, \ldots, d_u$. Let $e = e_1 + \cdots + e_t$. Now,

$$\operatorname{rank} P = f + e - t + s, \tag{19}$$

$$\operatorname{rank} Q = f + e - t + u, \tag{20}$$

$$\#(P/Q) = s, \tag{21}$$

$$\#(Q/P) = u, \tag{22}$$

$$m = f + e + s + p - t, \tag{23}$$

$$n = f + e + u + q - t, \tag{24}$$

$$\eta(P, Q) = t, \tag{25}$$

$$N(P, Q) = t - s - u, \tag{26}$$

$$L(P, Q) = \min(p, q) - t. \tag{27}$$

Let $\rho = \operatorname{rank} A$ and $\sigma = \operatorname{rank} B$. The inequalities (13–16) become

$$\rho \geqslant \operatorname{rank} P + \#(Q/P) = \operatorname{rank} Q + \#(P/Q) = f + e - t + s + u, \tag{13'}$$

$$\sigma \geqslant \operatorname{rank} P + \#(Q/P) = \operatorname{rank} Q + \#(P/Q) = f + e - t + s + u, \tag{14'}$$

$$\rho + \sigma \geqslant \operatorname{rank} P + \operatorname{rank} Q + \eta(P, Q) = 2(f + e) - t + u + s, \tag{15'}$$

$$\rho + \sigma \leqslant \min\{\text{rank } P + n, \text{rank } Q + m\} = 2(f + e - t) + s + u + \min(p, q). \tag{16'}$$

These inequalities define an isosceles trapezoid $T = T(P, Q)$ in the ρ, σ plane whose vertices are

$$(f + e - t + s + u,\ f + e) = (\text{rank } P + \#(Q/P),\ \text{rank } P + \#(Q/P) + N(P, Q)), \tag{28}$$

$$(f + e - t + s + u,\ f + e - t + \min(p, q)) = (\text{rank } P + \#(Q/P),\ \text{rank } P + \#(Q/P) + N(P, Q) + L(P, Q)), \tag{29}$$

$$(f + e, f + e - t + s + u) = (\text{rank } Q + \#(P/Q) + N(P, Q),\ \text{rank } Q + \#(P/Q)), \tag{30}$$

$$(f + e - t + \min(p, q), f + e - t + s + u) = (\text{rank } Q + \#(P/Q) + N(P, Q) + L(P, Q),\ \text{rank } Q + \#(P/Q)). \tag{31}$$

We now present our main result.

THEOREM 3. *The necessary and sufficient conditions that m-square matrix P and n-square matrix Q be representable as $P = AB$ and $Q = BA$ are that the elementary divisors of P and Q satisfy the condition given in Theorem* 1. *If this condition is satisfied then as A and B vary over all solutions of $P = AB$ and $Q = BA$, the integer pairs (ρ, σ) where $\rho =$* rank A *and $\sigma =$* rank B *precisely fill out the integer lattice points in the trapezoid $T(P, Q)$ for which the vertices are* (28), (29), (30), (31). *Equivalently, the conditions of Theorem* 1 *and the inequalities* (13′), (14′), (15′), (16′) *are necessary and sufficient in order that $P = AB$ and $Q = BA$ with* rank $A = \rho$ *and* rank $B = \sigma$.

Proof. Necessity is already established. Suppose $(\rho, \sigma) \in T$. First assume that $f + e \leqslant \rho \leqslant f + e - t + \min(p, q)$. Let $A = I_f \dotplus I_{\rho-f} \dotplus 0_{m-\rho, n-\rho}$. Let B be defined by (4). There $\bar{B}_{11}$ is to be f-square and is to have as elementary divisors the common set of elementary divisors of P and Q belonging to nonzero eigenvalues. We let $\tilde{B}_{11} = J_{e_1} \dotplus \cdots \dotplus J_{e_t} \dotplus 0_{\rho-e-f, \rho-e-f}$. Let $\tilde{B}_{12}$ be $(\rho - f) \times (m - \rho)$ with all columns zero except for the first s columns. The first s columns of $\tilde{B}_{12}$ each have a single nonzero entry, a one, and the one in column i of $\tilde{B}_{12}$ is opposite the last row (the row of zeros) of block $J_{e_{c_i}}$ in $\tilde{B}_{11}$; $1 \leqslant i \leqslant s$. We let $\tilde{B}_{21}$ be $(n-\rho) \times (\rho - f)$ with all rows zero except for the first u rows. Each of the first u rows contains a single nonzero entry, a one, and the one in row i of $\tilde{B}_{21}$ is underneath the first column (the column of zeros) of the block $J_{e_{d_i}}$ in $\tilde{B}_{11}$; $1 \leqslant i \leqslant u$. The matrix B_{22} is $(n-\rho) \times (m-\rho)$. In B_{22} each row opposite a one in $\tilde{B}_{21}$ is to be zero and each column beneath a one in $\tilde{B}_{12}$ is to be zero. This means that B_{22} is to be zero except for an $(n-\rho-u) \times (m-\rho-s)$ submatrix $\tilde{B}_{22}$ located in the lower right corner of B_{22}. The conditions on ρ ensure that $n-\rho-u \geqslant 0$ and $m-\rho-s \geqslant 0$. The proof of Theorem 1 now demonstrates that AB has the elementary divisors of P and that BA has the elementary divisors of Q, for any choice of $\tilde{B}_{22}$. The rank of B depends on the rank of $\tilde{B}_{22}$ and as rank $\tilde{B}_{22}$ varies over all possible values for an $(n-\rho-u) \times (m-\rho-s)$ matrix, rank B varies over all integers σ such that $f+e-t+s+u \leqslant \sigma \leqslant f+e-t+s+u+\min\{n-\rho-u, m-\rho-s\}$. Thus $\sigma = \operatorname{rank} B$ assumes any value such that $\sigma \geqslant f+e-t+s+u$ and $\rho+\sigma \leqslant 2(f+e-t)+s+u+\min(p,q)$. Thus for the given ρ, we achieve all σ permitted by (13′), (14′), (15′), (16′).

Next assume that $f+e-t+u+s \leqslant \rho < f+e$. For such a fixed ρ, we have to construct A and B with rank $A = \rho$ and rank $B = \sigma$, where the range of σ is determined by the inequalities (15′) and (16′). Observe that $t-s-u$ is the number of integers i, $1 \leqslant i \leqslant t$, distinct from all of $c_1, \ldots, c_s, d_1, \ldots, d_u$. By our choice of notation, each such i represents a coincident pair of nonlinear elementary divisors of P and Q; that is, $e_i \geqslant 2$. Since $0 < f+e-\rho \leqslant t-u-s$, from these coincident nonlinear pairs we may select $f+e-\rho$ elementary divisors and denote them by $\lambda^{e_1'}, \ldots, \lambda^{e'_{f+e-\rho}}$. Denote the remaining coincident nonlinear pairs of elementary divisors by $\lambda^{e_1''}, \ldots, \lambda^{e''_{t-u-s-f-e+\rho}}$. Now we construct A and B.

Let $A = I_f \dotplus I_{\rho-f} \dotplus 0_{m-\rho,n-\rho}$. Let B be given by (4) as before, where $\bar{B}_{11}$ is f-square and has as elementary divisors the common elementary divisors of P and Q belonging to nonzero eigenvalues. We let

$$\tilde{B}_{11} = J_{e_{c_1}} \dotplus \cdots \dotplus J_{e_{c_s}} \dotplus J_{e_{d_1}} \dotplus \cdots \dotplus J_{e_{d_u}} \dotplus J_{e_1'-1} \dotplus \cdots \dotplus J_{e'_{f+e-\rho}-1}$$
$$\dotplus J_{e_1''} \dotplus \cdots \dotplus J_{e''_{t-u-s-f-e+\rho}}.$$

Then $\tilde{B}_{11}$ is $(\rho - f)$-square. Matrix $\tilde{B}_{12}$ is $(\rho - f) \times (m - \rho)$ and is in Hermite column form. Each column of $\tilde{B}_{12}$ after the first $s + f + e - \rho$ columns is to be zero. Each of the first $s + f + e - \rho$ columns is to have a single nonzero entry, a one, the ones in these columns being opposite, respectively, the last rows (the zero rows) of the blocks $J_{e_{c_1}}, \ldots, J_{e_{c_s}}$, $J_{e_1'-1}, \ldots, J_{e'_{f+e-\rho}-1}$. The matrix $\tilde{B}_{21}$ is $(n - \rho) \times (\rho - f)$ and is in Hermite row form. Each row of $\tilde{B}_{21}$ after the first $u + f + e - \rho$ rows is to be zero. Each of the first $u + f + e - \rho$ rows is to have a single nonzero entry, a one, and the ones in these rows are, respectively, to be underneath the first columns (the columns of zeros) in $J_{e_{d_1}}, \ldots, J_{e_{d_u}}$, $J_{e_1'-1}, \ldots, J_{e'_{f+e-\rho}-1}$. The matrix B_{22} is to be $(n - \rho) \times (m - \rho)$. Each row of B_{22} opposite a one in $\tilde{B}_{21}$ is to be a zero row and each column of B_{22} beneath a one in $\tilde{B}_{12}$ is to be a zero column. Thus sitting in the lower right corner of B_{22} is an $(n - f - e - u) \times (m - f - e - s)$ matrix, call it $\tilde{B}_{22}$. For any choice of $\tilde{B}_{22}$, we find by the proof of Theorem 1 that AB has the elementary divisors of P and BA has the elementary divisors of Q. The rank of $\tilde{B}_{22}$ may assume any integer between zero and $\min(n - f - e - u, m - f - e - s) = \min(q - t, p - t)$. Thus rank B may assume any value σ for which $2(f + e) - t + u + s - \rho \leqslant \sigma \leqslant 2(f + e) - t + u + s - \rho + \min(q - t, p - t)$. That is, $\rho + \sigma \geqslant 2(f + e) - t + u + s$ and $\rho + \sigma \leqslant 2(f + e - t) + u + s + \min(p, q)$. This completes the proof of Theorem 3.

4. PROPERTIES OF THE FUNDAMENTAL TRAPEZOID

In this section we discuss properties of the trapezoid $T(P, Q)$, given matrices P and Q satisfying the necessary and sufficient conditions of Theorem 3. We continue the notation used above. We let $\mathscr{J}(a, b, c)$ denote the trapezoid with vertices (a, b), (a, c), (b, a), (c, a), when $0 \leqslant a \leqslant b \leqslant c$.

THEOREM 4. *$T(P, Q) = T(P_1, Q_1)$ if and only if*

$$\operatorname{rank} P + \#(Q/P) = \operatorname{rank} P_1 + \#(Q_1/P_1),$$

$$N(P, Q) = N(P_1, Q_1),$$

$$L(P, Q) = L(P_1, Q_1).$$

Proof. See (28–31).

We say $\mathscr{J}(a, b, c)$ is *admissible* if $\mathscr{J}(a, b, c) = T(P, Q)$ for some pair P, Q.

THEOREM 5. *$\mathscr{J}(a, b, c)$ is admissible if and only if $2a \geqslant b$.*

Proof. Let $\mathscr{J}(a, b, c) = T(P, Q)$. Then from (28–31) we get

$$\begin{aligned} f + e - t + s + u &= a, \\ f + e &= b, \\ f + e - t + \min(p, q) &= c. \end{aligned} \tag{32}$$

Hence $t - u - s = b - a$. Thus of the integers $e_1, \ldots, e_t$, exactly $b - a$ represent coincident nonlinear pairs of elementary divisors of P and Q for eigenvalue zero. For these coincident nonlinear pairs each $e_i \geqslant 2$. The remaining $u + s$ integers e_i are each $\geqslant 1$. Hence $e \geqslant 2(b - a) + u + s$. Since $f + e = a + t - s - u$, $f + 2(b - a) + u + s \leqslant a + (b - a)$, hence $f \leqslant 2a - b - u - s$. Since $f \geqslant 0$, we get $2a - b \geqslant 0$.

Conversely let $\mathscr{J}(a, b, c)$ be given with $2a \geqslant b$. We construct a pair P, Q with P similar to Q such that $\mathscr{J}(a, b, c) = T(P, Q)$. Set $u = s = 0$, $t = b - a$, $e_1 = \cdots = e_t = 2$, $p = q = c - a$, $f = 2a - b$. Then the required conditions $p \geqslant t$, $q \geqslant t$, $f \geqslant 0$ are satisfied, so that P and Q are constructible. Since $u = s = 0$ and $p = q$, in fact P and Q are similar. Then $f + e - t + s + u = a$, $f + e = b$, $f + e - t + \min(p, q) = c$, hence $T(P, Q) = \mathscr{J}(a, b, c)$, as required. Since independent similarities of P and Q may be made, we may achieve $P = Q$.

THEOREM 6. *Any admissible trapezoid $\mathscr{J}(a, b, c)$ is of the form $\mathscr{J}(a, b, c) = T(P, P)$.*

Theorem 6 leads us to seek additional properties that may be imposed on P. In the following theorems a point is considered to be a degenerate case of a triangle and a line is not considered to be a triangle.

THEOREM 7. *Let $\mathscr{J}(a, b, c)$ be admissible. Then $\mathscr{J}(a, b, c) = T(P, P)$ with P nilpotent and nonzero if and only if $\mathscr{J}(a, b, c)$ is not a triangle; that is, if and only if $b \neq a$. If $\mathscr{J}(a, b, c)$ is not a triangle then the set of permissible values for the index of nilpotency γ of P is precisely the set of integers γ for which $b/(b - a) \leqslant \gamma \leqslant 2a - b + 2$.*

We say P has index of nilpotency γ if $P^{\gamma-1} \neq 0$ but $P^{\gamma} = 0$.

Proof. Suppose $\mathscr{J}(a, b, c) = T(P, P)$ with P nilpotent and nonzero. Then $p = q$, $u = s = 0$, $f = 0$, and $t \neq 0$, hence $e - t = a$, $e = b$, $e - t + p = c$. Thus $b \neq a$, hence $\mathscr{J}$ is not a triangle. We have $e_1 \geqslant \cdots \geqslant e_t \geqslant 2$. The index of nilpotency γ of P is $\gamma = e_1$. From $e = b$ we get (using $t = b - a$) that $(b - a)\gamma \geqslant b$, hence $\gamma \geqslant b(b - a)^{-1}$. We also get from $e = b$ that $\gamma + 2(b - a - 1) \leqslant b$, hence $\gamma \leqslant 2a - b + 2$. This establishes the claimed bounds on γ.

Conversely suppose admissible nontriangular trapezoid $\mathscr{J}(a, b, c)$ is given and integer γ is given satisfying the bounds of the theorem. We now construct nilpotent P with γ as the index of nilpotency such that $\mathscr{J}(a, b, c) = T(P, P)$. Observe that $b(b - a)^{-1} > 1$ since $b(b - a)^{-1} \leqslant 1$ implies $a = 0$, hence $b = 0$ (because $2a \geqslant b$), and hence $a = b$, a contradiction. Thus $\gamma \geqslant 2$. We now attempt to satisfy the equations $f = u = s = 0$, $p = q$, $e_1 = \gamma$, $e - t = a$, $e = b$, $e - t + p = c$. We take $t = b - a$, $p = c - a$. We have only to choose $e_1 = \gamma, e_2, \ldots, e_{b-a}$ such that $\gamma \geqslant e_2 \geqslant \cdots \geqslant e_{b-a} \geqslant 2$ and such that $e = b$. As $e_2, \ldots, e_{b-a}$ range independently over the integers between 2 and γ, $e = \gamma + e_2 + \cdots + e_{b-a}$ assumes all integral values between $\gamma + 2(b - a - 1)$ and $(b - a)\gamma$. Since $\gamma + 2(b - a - 1) \leqslant b$ and $(b - a)\gamma \geqslant b$, there is a choice of $e_2, \ldots, e_{b-a}$, each between 2 and γ, such that $e = b$. This completes the proof.

THEOREM 8. *Let $\mathscr{J}(a, b, c)$ be admissible. As P ranges over all matrices such that $\mathscr{J}(a, b, c) = T(P, P)$ the number of nonzero eigenvalues (counting multiplicities) that P may have is precisely the integers f for which $0 \leqslant f \leqslant 2a - b$.*

Proof. Using $p = q$ and $u = s = 0$, we observed in the proof of Theorem 5 that $f \leqslant 2a - b$. This gives the bound of the theorem. Conversely, given f, we have to solve $f + e - t = a$, $f + e = b$, $f + e - t + p = c$. Put $t = b - a$, $e_1 = b - f - 2(b - a - 1)$, $e_2 = \cdots = e_t = 2$, $p = c - a$. Then $e_1 \geqslant 2$ as required since $f \leqslant 2a - b$.

THEOREM 9. *$T(P, Q)$ is a point if and only if $N(P, Q) = L(P, Q) = 0$.*

Proof. Let $T(P, Q) = \mathscr{J}(a, b, c)$. $\mathscr{J}(a, b, c)$ is a point if and only if $a = b = c$ and from (28–31) this happens if and only if $N(P, Q) = L(P, Q) = 0$.

COROLLARY. *In* $P = AB$, $Q = BA$ *the integers* rank A, rank B *are uniquely determined by* P *and* Q *if and only if* $N(P, Q) = L(P, Q) = 0$. *In this event* rank A = rank B = rank $P + \#(Q/P)$ = rank $Q + \#(P/Q)$.

COROLLARY. $T(P, P)$ *is a point if and only if* P *is nonsingular.*

THEOREM 10. $T(P, Q)$ *is a line if and only if* $L(P, Q) = 0$.

Proof. $\mathscr{J}(a, b, c)$ is a line if and only if $b = c$. By (28–31) this happens if and only if $L(P, Q) = 0$.

COROLLARY. *In* $P = AB$, $Q = BA$, *the sum* rank A + rank B *is uniquely determined by* P *and* Q *if and only if* $L(P, Q) = 0$. *In this case the constant value of the sum is* rank P = rank $Q + \eta(P, Q)$, *for* rank $P + \#(Q/P) \leqslant$ rank A, rank $B \leqslant$ rank $P + \#(Q/P) + N(P, Q)$.

THEOREM 11. $T(P, Q)$ *is a triangle if and only if* $N(P, Q) = 0$.

Proof. $\mathscr{J}(a, b, c)$ is a triangle if and only if $b = a$. Using (28–31) the result follows.

COROLLARY. $T(P, P)$ *is a triangle if and only if* P *has only linear elementary divisors for eigenvalue zero.*

COROLLARY. *Triangular trapezoids, and only triangular trapezoids, are of the form* $T(P, P)$ *for diagonable* P.

THEOREM 12. *Let* $\mathscr{J}(a, b, c)$ *be admissible. Then for any* P, Q *such that* $\mathscr{J}(a, b, c) = T(P, Q)$ *we have:*

(i) $N(P, Q) = b - a$;
(ii) $L(P, Q) = c - a$;
(iii) degree $P \geqslant c$, degree $Q \geqslant c$ *with* degree P = degree $Q = c$ *if and only if* P *and* Q *are similar*;
(iv) rank $P \leqslant a$, rank $Q \leqslant a$ *with* rank $P = a$ *if and only if* $\#(Q/P) = 0$ *and* rank $Q = a$ *if and only if* $\#(P/Q) = 0$.

In particular, when degree P = degree Q, rank P = rank $Q = a$ *if and only if* P *and* Q *are similar.*

Proof. (i) and (ii) follow from (28–31). For (iii) observe that degree $P = f + e + s + p - t \geqslant f + e - t + \min(p, q) = c$ and degree $Q = f + e + u + q - t \geqslant f + e - t + \min(p, q) = c$, with both inequalities being equality if and only if $s = u = 0$ and $p = q$. This proves (iii). For (iv) observe that rank $P = f + e - t + s \leqslant f + e - t + s + u = a$ and rank $Q = f + e - t + u \leqslant f + e - t + s + u = a$.

THEOREM 13. *The number of admissible line trapezoids $T(P, P)$, where* rank $P = a$, *is exactly $a + 1$. These different admissible trapezoids arise as $N(P, P)$ assumes the values $0, 1, \ldots, a$; $L(P, P) = 0$; and as* degree P *assumes the values $a, a + 1, \ldots, 2a$.*

Proof. $\mathscr{J}(a, b, c)$ is admissible if and only if $2a \geqslant c$. Hence c can only assume the values $a, a + 1, \ldots, 2a$.

THEOREM 14. *The number of integral pairs $(\rho, \sigma) \in T(P, Q)$ is*

$$\tfrac{1}{2}(1 + L(P, Q))(2 + 2N(P, Q) + L(P, Q)).$$

Proof. Direct computation.

REFERENCES

1 H. Flanders, Elementary divisors of AB and BA, *Proc. Amer. Math. Soc.* **2**(1951), 871–874.
2 W. V. Parker and B. E. Mitchell, Elementary divisors of certain matrices, *Duke Math. J.* **19**(1952), 483–485.

Received July 14, 1967