SIMILARITY OF BLOCK DIAGONAL AND BLOCK TRIANGULAR MATRICES

Vadim Olshevsky

Integral Equations and Operator Theory, 15 (1992), 853-863.

It is shown that if a block triangular matrix is similar to its block diagonal part, then the similarity matrix can be chosen of the block triangular form. An analogous statement is proved for equivalent matrices. For the simplest case of 2×2 block matrices these results were obtained by W.Roth [1]. It is shown that all these results do not admit a generalization for the infinite dimensional case.

0. INTRODUCTION

Let $A \in \mathrm{C}^{n\times n}$, $B \in \mathrm{C}^{m\times m}$, $C \in \mathrm{C}^{n\times m}$, where $\mathrm{C}^{r\times s}$ stands for the set of all $r \times s$ matrices with complex entries. Consider the matrix equations

$$AX - XB = C, \tag{0.1}$$

and

$$AX - YB = C, \tag{0.2}$$

where X and Y are unknown matrices from $\mathrm{C}^{n\times m}$. It is easy to see that equation (0.1) is solvable if and only if there exists a matrix $X \in \mathrm{C}^{n\times m}$ fulfilling the following matrix identity:

$$\begin{pmatrix} I & -X \\ 0 & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}. \tag{0.3}$$

Analogously, the solvability of equation (0.2) is equivalent to the existence of matrices X and Y from $\mathrm{C}^{n\times m}$ such that

$$\begin{pmatrix} I & -Y \\ 0 & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}. \tag{0.4}$$

Note that the first factor in the left hand side of (0.3) is the inverse of the third factor and therefore from (0.3) there follows that if equation (0.1) is solvable then the matrices

$$E = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \quad \text{and} \quad F = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \tag{0.5}$$

are similar. Analogously, from (0.4) it follows that the solvability of equation (0.2) implies the equivalence of the matrices E and F from (0.5) (equal rank).

W.Roth [1] obtained the converse for the last two statements, namely: if matrices E and F from (0.5) are similar, then equation (0.1) has a solution. Moreover, it was proved in [1] that the equivalence of matrices E and F implies the solvability of (0.2). These results imply that if matrices E and F from (0.5) are similar then the similarity matrix can be chosen to be of the upper block triangular form $\begin{pmatrix} I & X \\ 0 & I \end{pmatrix}$. Analogously, if E and F are equivalent, then one can choose the equivalence matrices to be of the upper block triangular forms $\begin{pmatrix} I & -Y \\ 0 & I \end{pmatrix}$ and $\begin{pmatrix} I & X \\ 0 & I \end{pmatrix}$.

Since [1] a number of different proofs of Roth's theorems were obtained mostly passing to the corresponding canonical forms (see for example [2-5]). In section 1 of the present paper we give short new proofs based on geometrical arguments.

In section 2 it is shown that Roth's theorems do not admit a generalization for infinite dimensional spaces (compare with [6,p.17]). Here we shall give an example of operators A, B and C acting in an infinite dimensional separable Hilbert space for which the corresponding operators E and F from (0.5) are similar, but nevertheless equations (0.1) and (0.2) have no solution.

In the third section we investigate the generalizations of both Roth's theorems for block matrices of arbitrary order. It is proved here that if an upper block triangular matrix is similar to its block diagonal part, then the similarity matrix can be chosen to be of an upper block triangular form. More precisely, let $A_{ij} \in \mathbb{C}^{n_i \times n_j}$ $(1 \le i \le j \le k)$. It is shown that if the block matrices

$$G = \begin{pmatrix} A_{11} & 0 & \cdots & \cdots & 0 \\ 0 & A_{22} & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & A_{kk} \end{pmatrix} \quad \text{and } H = \begin{pmatrix} A_{11} & A_{12} & \cdots & \cdots & A_{1k} \\ 0 & A_{22} & A_{23} & \cdots & A_{2k} \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & A_{kk} \end{pmatrix} \tag{0.6}$$

are similar, then there exists matrices $X_{ij} \in \mathbb{C}^{n_i \times n_j}$ $(1 \le i < j \le k)$ such that the following identity holds :

$$\begin{pmatrix} I_1 & X_{12} & \cdots & \cdots & X_{1k} \\ 0 & I_2 & X_{23} & \cdots & X_{2k} \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & I_k \end{pmatrix}^{-1} \cdot G \cdot \begin{pmatrix} I_1 & X_{12} & \cdots & \cdots & X_{1k} \\ 0 & I_2 & X_{23} & \cdots & X_{2k} \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & I_k \end{pmatrix} = H.$$

It is also proved, that if G and H in (0.6) are equivalent then there exist matrices X_{ij}, Y_{ij} $(1 \le i < j \le k)$ such that

$$\begin{pmatrix} I_1 & Y_{12} & \cdots & \cdots & Y_{1k} \\ 0 & I_2 & Y_{23} & \cdots & Y_{2k} \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & I_k \end{pmatrix} \cdot G \cdot \begin{pmatrix} I_1 & X_{12} & \cdots & \cdots & X_{1k} \\ 0 & I_2 & X_{23} & \cdots & X_{2k} \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & I_k \end{pmatrix} = H.$$

The proof in the equivalence case is simpler than the proof in the similarity case, although both are carried out by the same arguments. Note that the assertion concerning the equivalence can be deduced also from theorem 2 in [3].

The author wishes to thank I.Gohberg and A.Markus for fruitful discussions, and N.Cohen, who carefully read the manuscript and made a number of helpful comments.

1. 2×2 BLOCK MATRICES

THEOREM 1.1 *Let $A \in \mathbb{C}^{n\times n}, B \in \mathbb{C}^{m\times m}$ and $C \in \mathbb{C}^{n\times m}$. The following conditions are equivalent.*

(i). *The equation $AX - XB = C$ has a solution $X \in \mathbb{C}^{n\times m}$.*

(ii). *There exists an $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ -invariant subspace $W \subset \mathbb{C}^{n+m}$ such that $U\dot{+}W = \mathbb{C}^{n+m}$, where U is the subspace of all vectors from $\mathbb{C}^{n+m}$ whose last m coordinates are equal to zero.*

(iii). *The matrices $E = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ and $F = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ are similar.*

(iv). *There exists $X \in \mathbb{C}^{n\times m}$ such that*

$$\begin{pmatrix} I & -X \\ 0 & I \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} I & X \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}.$$

THEOREM 1.2 *Let $A \in \mathbb{C}^{n\times n}, B \in \mathbb{C}^{m\times m}$ and $C \in \mathbb{C}^{n\times m}$. The following conditions are equivalent.*

(i). *The equation $AX - YB = C$ has a solution $X, Y \in \mathbb{C}^{n\times m}$.*

(ii). $C(\mathrm{Ker}\ B) \subset \mathrm{Im}\ A$.

(iii). *The matrices $E = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ and $F = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ are equivalent.*

(iv). *There exists $X,\ Y \in \mathbb{C}^{n\times m}$ such that*

$$\begin{pmatrix} I & -Y \\ 0 & I \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} I & X \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}.$$

Note that the equivalence of conditions (i) and (iii) in both theorems is the assertion of Roth's theorems. As was mentioned, conditions (iv) are evidently equivalent to the corresponding conditions (i) and are included in the formulation for the convenience of the proofs. The equivalence of the conditions (i) and (ii) in theorem 1.1 is well known in the more general framework of the theory of matrix Ricatti equations (see, for example, [7, p. 545]). The assertion about the equivalence of condition (ii) in theorem 1.2 to the other conditions seems to be new.

PROOF OF THEOREM 1.1. The proof of the implications (i) $\Rightarrow$ (iv)$\Rightarrow$ (iii) poses no difficulties. Assume therefore that condition (iii) holds and prove (ii). Alongside with the matrices E and F consider also the matrix

$$F_\alpha = \begin{pmatrix} A & \alpha C \\ 0 & B \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & \alpha I \end{pmatrix}^{-1} F \begin{pmatrix} I & 0 \\ 0 & \alpha I \end{pmatrix} \qquad (\alpha \in \mathrm{C}). \tag{1.1}$$

In [7, p.488] it was shown that for every matrix E there is a constant $K > 0$ such that for every matrix H which is similar to E we have

$$\|I - S\| \leq K\|H - E\|$$

for some nonsingular matrix S, satisfying $S^{-1}HS = E$. Since for every complex α the matrix F_α from (1.1) is similar to E, one can choose a matrix S_α such that

$$S_\alpha^{-1}F_\alpha S_\alpha = E, \quad \|I - S_\alpha\| \leq K\|F_\alpha - E\|. \tag{1.2}$$

Take a number α so small, that the right hand side of inequality (1.2) should be less than 1. Denote by V the subspace of all vectors from C^{n+m} whose first n coordinates are equal to zero. Obviously, the subspace $S_\alpha(V)$ is F_α-invariant and from the condition $\|I - S_\alpha\| < 1$ it follows that

$$U \dot{+} S_\alpha(V) = \mathrm{C}^{n+m}. \tag{1.3}$$

Put $W = \begin{pmatrix} I & 0 \\ 0 & \alpha I \end{pmatrix} S_\alpha(V)$. From (1.1) and (1.2) it follows that the subspace $W \subset \mathrm{C}^{n+m}$ satisfies all the claims of condition (ii).

To complete the proof it remains only to show that condition (ii) implies (i). Denote by S any nonsingular matrix which maps U onto W. Evidently, with respect to the decomposition of C^{n+m} as the direct sum of U and V, the matrix S has the following block triangular form: $S = \begin{pmatrix} I & Z \\ 0 & Y \end{pmatrix}$, where $Z \in \mathrm{C}^{n\times m}$ and Y is a nonsingular matrix from $\mathrm{C}^{m\times m}$. Since W is F-invariant, it follows that $V = S^{-1}(W)$ is $S^{-1}FS$-invariant and therefore the block $AZ - ZY^{-1}BY + CY$ in the upper right corner of the matrix $S^{-1}FS$ is identically equal to zero. It means that the matrix $X = -ZY^{-1}$ is a solution of the equation $AX - XB = C$. ■

PROOF OF THEOREM 1.2. As in the proof of the previous theorem, the implications (i)$\Rightarrow$ (iv) and (iv)$\Rightarrow$ (iii) are obvious, therefore let us turn to the proof of (iii)$\Rightarrow$ (ii). It is easy to see that

$$\mathrm{Ker}\ E = \mathrm{Ker} \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} \dot{+} \mathrm{Ker} \begin{pmatrix} I & 0 \\ 0 & B \end{pmatrix}.$$

Denote by $W \subset \mathrm{C}^{n+m}$ some direct complement to the subspace $\mathrm{Ker} \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix}$ in $\mathrm{Ker}\ F$, i.e.

$$\mathrm{Ker}\ F = \mathrm{Ker} \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} \dot{+} W.$$

Matrices E and F are equivalent if and only if they have equal rank, or, in other words, when $\dim W = \dim \operatorname{Ker} \begin{pmatrix} I & 0 \\ 0 & B \end{pmatrix}$. The latter equality is satisfied only if the condition (ii) holds, since $\begin{pmatrix} x \\ y \end{pmatrix} \in W$ means that $y \in \operatorname{Ker} B$, and $Cy = -Ax$.

To complete the proof it remains to show that (ii) implies (i). Denote by W the subspace of all vectors with zero first $m-k$ coordinates. Here $k = \dim \operatorname{Ker} B$. Let R be any nonsingular matrix from $\mathrm{C}^{m\times m}$, which maps W onto Ker B. Put $\hat{B} = BR$, $\hat{C} = CR$ and consider the equation

$$AZ - Y\hat{B} = \hat{C}. \tag{1.4}$$

Write $\hat{B}$, $\hat{C}$ and Z in the form

$$\hat{B} = [B_1\ 0], \quad \hat{C} = [C_1\ C_2] \ \text{ and } \ Z = [Z_1\ Z_2],$$

where $B_1 \in \mathrm{C}^{m\times(m-k)}$, $C_1,\ Z_1 \in \mathrm{C}^{n\times(m-k)}$, $C_2,\ Z_2 \in \mathrm{C}^{n\times k}$. Note that

$$\operatorname{Ker} B_1 = \{0\}, \tag{1.5}$$

and that condition (ii) yields

$$\operatorname{Im} C_2 \subset \operatorname{Im} A. \tag{1.6}$$

Equation (1.4) is equivalent to the following system of equations:

$$AZ_1 - YB_1 = C_1 \ \text{ and } \ AZ_2 = C_2.$$

Condition (1.5) implies that the first of these equations is solvable with respect to Y for every choice of Z_1. The second equation is solvable with respect to Z_2 due to (1.6). From the last arguments it follows that equation (1.4) has a solution Z, Y. It remains only to note, that in this case the matrices $X = ZR^{-1}$ and Y satisfy the equation $AX - YB = C$. ■

2. COUNTEREXAMPLE IN THE INFINITE DIMENSIONAL CASE

Here we shall give an example showing that in infinite dimensional space the similarity of the operators

$$E = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \ \text{ and } \ F = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \tag{2.1}$$

no longer implies the solvability of the operator equation $AX - XB = C$. The same example also shows that the equivalence of the operators E and F is not sufficient for the solvability of the equation $AX - YB = C$, and so, in this case, both Roth's theorems are not valid.

Let H be an infinite dimensional separable Hilbert space with orthonormal basis $\{\mathbf{e}_i\}_1^\infty$. Let the operators A and C, acting on H be defined as follows:

$$A\mathbf{e}_{3k+1} = \mathbf{0}, \quad A\mathbf{e}_{3k+2} = \mathbf{0}, \quad A\mathbf{e}_{3k+3} = \mathbf{e}_{3k+2} \qquad (k = 0, 1, 2, ...),$$

$$C\mathbf{e}_1 = \mathbf{e}_1, \quad \text{and} \quad C\mathbf{e}_i = \mathbf{0} \ for \ \ i \neq 1.$$

Let $B = A$. It is easy to see, that operator E from (2.1) has the only eigenvalue $\lambda_0 = 0$ and that to this eigenvalue there corresponds a countable number of Jordan chains of length 1 and a countable number of Jordan chains of length 2. Moreover, the vectors of these chains form the orthonormal basis of $H \oplus H$. Since operator F from (2.1) has the same properties, therefore these operators are equivalent and, moreover, similar. On the other hand, the operator equation $AX - YB = C$ (and, what is more, the equation $AX - XB = C$) is not solvable. Indeed, the assumption that solutions X, Y exist would imply that equality $AX\mathbf{e}_1 = \mathbf{e}_1$ holds, which is a contradiction since $\mathbf{e}_1 \notin \operatorname{Im} A$.

3. BLOCK MATRICES OF ARBITRARY ORDER

In this section we will prove the following two statements.

THEOREM 3.1 *Let* $A_{ij} \in \mathbb{C}^{n_i \times n_j}$ $(1 \leq i \leq j \leq k)$. *Assume that the matrices*

$$G = \begin{pmatrix} A_{11} & 0 & \cdots & \cdots & 0 \\ 0 & A_{22} & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & A_{kk} \end{pmatrix} \quad \text{and} \quad H = \begin{pmatrix} A_{11} & A_{12} & \cdots & \cdots & A_{1k} \\ 0 & A_{22} & A_{23} & \cdots & A_{2k} \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & A_{kk} \end{pmatrix} \tag{3.1}$$

are similar. Then there exist matrices $X_{ij} \in \mathbb{C}^{n_i \times n_j}$ $(1 \leq i < j \leq k)$, *such that*

$$\begin{pmatrix} I_1 & X_{12} & \cdots & \cdots & X_{1k} \\ 0 & I_2 & X_{23} & \cdots & X_{2k} \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & I_k \end{pmatrix}^{-1} \cdot G \cdot \begin{pmatrix} I_1 & X_{12} & \cdots & \cdots & X_{1k} \\ 0 & I_2 & X_{23} & \cdots & X_{2k} \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & I_k \end{pmatrix} = H.$$

THEOREM 3.2 *Let* $A_{ij} \in \mathbb{C}^{n_i \times n_j}$ $(1 \leq i \leq j \leq k)$ *be such that block matrices* G *and* H *from* (3.1) *are equivalent. Then there exist matrices* $X_{ij}, Y_{ij} \in \mathbb{C}^{n_i \times n_j}$ $(1 \leq i < j \leq k)$, *such that*

$$\begin{pmatrix} I_1 & Y_{12} & \cdots & \cdots & Y_{1k} \\ 0 & I_2 & Y_{23} & \cdots & Y_{2k} \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & I_k \end{pmatrix} \cdot G \cdot \begin{pmatrix} I_1 & X_{12} & \cdots & \cdots & X_{1k} \\ 0 & I_2 & X_{23} & \cdots & X_{2k} \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & I_k \end{pmatrix} = H.$$

For the proof of theorem 3.1 we shall need the following auxiliary proposition.

LEMMA 3.3 *Let R and S be matrices from $\mathrm{C}^{n\times n}$. If there exist matrices S_i ($i = 1,2,...$), which are similar to S, but converge to R, and there exist matrices R_i ($i = 1,2,...$), which are similar to R, but converge to S, then matrices S and R are similar.*

One can deduce lemma 3.3 from [8,9] (see also [7, p.476]), but we prefer to give a direct and simple proof.

PROOF. The conditions of the lemma imply, that for every complex λ and natural k

$$(R-\lambda I)^k = \lim_{i\to\infty}(S_i-\lambda I)^k \ , \quad (S-\lambda I)^k = \lim_{i\to\infty}(R_i-\lambda I)^k \ .$$

Since passage to the limit cannot decrease the dimension of the kernel of a matrix, it follows from the last equalities that

$$\dim \mathrm{Ker}\ (R-\lambda I)^k = \dim \mathrm{Ker}\ (S-\lambda I)^k \quad (k=1,2,...\ \ ;\ \ \lambda\in \mathrm{C}).$$

These equalities can hold only in case all the eigenvalues of matrices R and S as well as the corresponding sizes of Jordan cells coincide, or in other words when R and S are similar. ■

PROOF OF THEOREM 3.1. Set $D_0 = H$, $D_k = G$ and consider the matrices

$$D_l = \left(\begin{array}{cccc|cccc} A_{11} & 0 & \cdots & 0 & & & & \\ 0 & A_{22} & \ddots & \vdots & & 0 & & \\ \vdots & \ddots & \ddots & 0 & & & & \\ 0 & \cdots & 0 & A_{ll} & & & & \\ \hline & & & & A_{l+1,l+1} & \cdots & \cdots & A_{l+1,k} \\ & 0 & & & 0 & \ddots & & \vdots \\ & & & & \vdots & \ddots & \ddots & \vdots \\ & & & & 0 & \cdots & 0 & A_{kk} \end{array}\right) \qquad (l=1,2,...,k-1), \qquad (3.2)$$

which are derived from the matrix H by annihilating all the entries of the first l rows which are to the right of the main diagonal. Let us show that all the matrices D_l ($l = 0,1,...,k$) are similar. Consider for $n \in$ the following block diagonal matrices:

$$Q_{ln} = \mathrm{diag}\{I_1,\ \frac{1}{n}I_2,\ \frac{1}{n^2}I_3,\ ...\ ,\ \frac{1}{n^{l-1}}I_l,\ \frac{1}{n^l}I_{l+1},\ \frac{1}{n^l}I_{l+2},\ ...\ ,\ \frac{1}{n^l}I_k\},$$

and

$$R_n = \mathrm{diag}\{I_1,\ \frac{1}{n}I_2,\ \frac{1}{n^2}I_3,\ ...\ ,\ \frac{1}{n^{k-1}}I_k\}.$$

It is easy to see that

$$\lim_{n\to\infty}\|Q_{ln}^{-1}HQ_{ln}-D_l\| = 0; \quad \lim_{n\to\infty}\|R_n^{-1}D_lR_n-G\| = 0. \qquad (3.3)$$

According to the lemma 3.3 all the matrices D_l ($l = 1,2,...,k-1$) and G are similar. From theorem 1.1 it follows that for

$$D_{l-1} = \left(\begin{array}{cccc|cccc} A_{11} & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & A_{22} & \ddots & \vdots & \vdots & & & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & A_{ll} & A_{l,l+1} & \cdots & \cdots & A_{lk} \\ \hline & & & & A_{l+1,l+1} & \cdots & \cdots & A_{l+1,k} \\ & \multicolumn{2}{c}{\mathbf{0}} & & 0 & \ddots & & \vdots \\ & & & & \vdots & \ddots & \ddots & \vdots \\ & & & & 0 & \cdots & 0 & A_{kk} \end{array}\right),$$

$$D_l = \left(\begin{array}{cccc|cccc} A_{11} & 0 & \cdots & 0 & & & & \\ 0 & \ddots & \ddots & \vdots & & \multicolumn{2}{c}{\mathbf{0}} & \\ \vdots & \ddots & \ddots & 0 & & & & \\ 0 & \cdots & 0 & A_{ll} & & & & \\ \hline & & & & A_{l+1,l+1} & \cdots & \cdots & A_{l+1,k} \\ & \multicolumn{2}{c}{\mathbf{0}} & & 0 & \ddots & & \vdots \\ & & & & \vdots & \ddots & \ddots & \vdots \\ & & & & 0 & \cdots & 0 & A_{kk} \end{array}\right)$$

there exists matrix X_l, such that $T_l^{-1} D_{l-1} T_l = D_l$, where

$$T_l = \left(\begin{array}{cccc|cccc} I_1 & 0 & \cdots & 0 & & & & \\ 0 & \ddots & \ddots & \vdots & & \multicolumn{2}{c}{X_l} & \\ \vdots & \ddots & \ddots & 0 & & & & \\ 0 & \cdots & 0 & I_l & & & & \\ \hline & & & & I_{l+1} & 0 & \cdots & 0 \\ & \multicolumn{2}{c}{\mathbf{0}} & & 0 & \ddots & \ddots & \vdots \\ & & & & \vdots & \ddots & \ddots & 0 \\ & & & & 0 & \cdots & 0 & I_k \end{array}\right).$$

Put $T = \prod_{i=1}^{k} T_i$. Obviously, $T^{-1}GT = H$ and T has the following form:

$$T = \begin{pmatrix} I_1 & X_{12} & \cdots & \cdots & X_{1k} \\ 0 & I_2 & X_{23} & \cdots & X_{2k} \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & I_k \end{pmatrix}.$$

■

PROOF OF THEOREM 3.2. As in the proof of the previous theorem, consider the matrices D_l from (3.2) and show that all of them are equivalent. Since passage to the limit cannot increase the rank, it follows from (3.3) that if rank $D_l <$ rank H for some l, then rank $G <$ rank H. But the last inequality contradicts the equivalence of G and H. So, all of D_l have equal rank and hence are equivalent. According to theorem 3.2 for every l $(1 \le l \le k)$ there exist matrices X_l, Y_l of corresponding sizes, such that $S_l D_{l-1} T_l = D_l$, where

$$T_l = \left(\begin{array}{cccc|cccc} I_1 & 0 & \cdots & 0 & & & & \\ 0 & \ddots & \ddots & \vdots & & X_l & & \\ \vdots & \ddots & \ddots & 0 & & & & \\ 0 & \cdots & 0 & I_l & & & & \\ \hline & & & & I_{l+1} & 0 & \cdots & 0 \\ & & 0 & & 0 & \ddots & \ddots & \vdots \\ & & & & \vdots & \ddots & \ddots & 0 \\ & & & & 0 & \cdots & 0 & I_k \end{array}\right),$$

$$S_l = \left(\begin{array}{cccc|cccc} I_1 & 0 & \cdots & 0 & & & & \\ 0 & \ddots & \ddots & \vdots & & Y_l & & \\ \vdots & \ddots & \ddots & 0 & & & & \\ 0 & \cdots & 0 & I_l & & & & \\ \hline & & & & I_{l+1} & 0 & \cdots & 0 \\ & & 0 & & 0 & \ddots & \ddots & \vdots \\ & & & & \vdots & \ddots & \ddots & 0 \\ & & & & 0 & \cdots & 0 & I_k \end{array}\right).$$

Put $S = \prod_{i=1}^{k} S_i$, $T = \prod_{i=1}^{k} T_i$. Evidently, $SGT = H$ and the matrices S and T have the following desired form:

$$T = \begin{pmatrix} I_1 & X_{12} & \cdots & \cdots & X_{1k} \\ 0 & I_2 & X_{23} & \cdots & X_{2k} \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & I_k \end{pmatrix}, \quad S = \begin{pmatrix} I_1 & Y_{12} & \cdots & \cdots & Y_{1k} \\ 0 & I_2 & Y_{23} & \cdots & Y_{2k} \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & I_k \end{pmatrix}.$$

$\blacksquare$

As was noted in the introduction, theorem 3.2 can be deduced also from theorem 2 in [3].

REFERENCES

[**1**] W.E.Roth,*The equation $AX - YB = C$ and $AX - XB = C$ in matrices*, Proc. Amer. Math. Soc., **3** (1952), 392 - 396.

[**2**] R.Feinberg, *Similarity of partitioned matrices*, J. of Res. Nat. Bur. of Standards. B., **79B** (1975), 117 - 125.

[**3**] R.Feinberg, *Equivalence of partitioned matrices*, J. of Res. Nat. Bur. of Standards. B., **80B** (1976), 89 - 97.

[**4**] H.Flanders and H.Wimmer, *On the matrix equations $AX - XB = C$ and $AX - YB = C$*, Siam J. Appl. Math., **32**, No. 4 (1977), 707 -710.

[**5**] R.Horn and C.Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1991.

[**6**] I.Gohberg, S.Goldberg and M.Kaashoek, *Classes of Linear Operators Vol.I*, Birkhauser Verlag, Basel, 1990.

[**7**] I.Gohberg, P.Lancaster and L.Rodman, *Invariant Subspaces of Matrices*, J.Wiley, New York, 1986.

[**8**] A.Markus and E.Parilis, *The change of the Jordan structure of a matrix under small perturbations*, Linear Algebra Appl., **54** (1983), 139 - 152.

[**9**] H. den Boer and G.P.Thijsse, *Semistability of sums of partial multiplicities under small perturbation*, Integral Equations and Operator Theory, **3** (1980), 23 - 42.

School of Mathematical Sciences
Raymond and Beverly Sackler
Faculty of Exact Sciences
Tel Aviv University
Ramat Aviv 69978, ISRAEL

MSC 1991 : Primary 15A24, Secondary 47A62

Submitted : November 3, 1991