ISRAEL JOURNAL OF MATHEMATICS, Vol. 35, Nos. 1-2, 1980

ON POINTWISE AND ANALYTIC SIMILARITY OF MATRICES

BY

SHMUEL FRIEDLAND†

ABSTRACT

Let $A(\varepsilon)$ and $B(\varepsilon)$ be complex valued matrices analytic in ε at the origin. $A(\varepsilon)\sim_p B(\varepsilon)$ if $A(\varepsilon)$ is similar to $B(\varepsilon)$ for any $|\varepsilon|<r$, $A(\varepsilon)\sim_a B(\varepsilon)$ if $B(\varepsilon)=T(\varepsilon)A(\varepsilon)T^{-1}(\varepsilon)$ and $T(\varepsilon)$ is analytic and $|T(\varepsilon)|\neq 0$ for $|\varepsilon|<r$! In this paper we find a necessary and sufficient conditions on $A(\varepsilon)$ and $B(\varepsilon)$ such that $A(\varepsilon)\sim_a B(\varepsilon)$ provided that $A(\varepsilon)\sim_p B(\varepsilon)$. This problem arises in study of certain ordinary differential equations singular with respect to a parameter ε in the origin and was first stated by Wasow.

1. Introduction

Let $A(\varepsilon)$ and $B(\varepsilon)$ be $n\times n$ complex valued matrices analytic in a parameter ε, in $D_r=\{z,|z|<r\}$ for some $r>0$. We call such matrices analytic at the origin. That is we have the Mclaurin expansions

$$A(\varepsilon)=\sum_{k=0}^{\infty}A_k\varepsilon^k,\quad B(\varepsilon)=\sum_{k=0}^{\infty}B_k\varepsilon^k,\qquad A_k,B_k\in M_n(\mathbf{C}) \tag{1.1}$$

which converge in D_r. One says that $A(\varepsilon)$ and $B(\varepsilon)$ are pointwise similar in D_r (denote it by $A(\varepsilon)\sim_p B(\varepsilon)$) if $A(\varepsilon)$ and $B(\varepsilon)$ are similar for any $\varepsilon\in D_r$. $A(\varepsilon)$ and $B(\varepsilon)$ are said to be analytically similar in $D_{r'}$ (denote it by $A(\varepsilon)\sim_a B(\varepsilon)$) if there exists $T(\varepsilon)$,

$$T(\varepsilon)=\sum_{k=0}^{\infty}T_k\varepsilon^k,\qquad T_k\in M_n(\mathbf{C})\quad (\text{convergent for } |\varepsilon|<r'), \tag{1.2}$$

such that

$$|T(\varepsilon)|\neq 0\qquad \text{for } |\varepsilon|<r' \tag{1.3}$$

(here by $|T|$ we denote the determinant of T) and

† Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

Received March 16, 1979

$$B(\varepsilon) = T(\varepsilon)A(\varepsilon)T^{-1}(\varepsilon). \tag{1.4}$$

The problem of determining whether two given analytic valued matrices $A(\varepsilon)$ and $B(\varepsilon)$ are analytically similar in $D_{r'}$ for some $r' > 0$ is important in study of certain ordinary differential equations singular with respect to a parameter ε in the origin (e.g. see [4] and references therein). Clearly if $A(\varepsilon) \sim_a B(\varepsilon)$ in $D_{r'}$ then $A(\varepsilon) \sim_p B(\varepsilon)$ in $D_{r'}$. Naturally one poses the following question:

PROBLEM 1.1. (Wasow [4]) Assume that $A(\varepsilon) \sim_p B(\varepsilon)$ in D_r. What other conditions should $A(\varepsilon)$ and $B(\varepsilon)$ satisfy in order that $A(\varepsilon) \sim_a B(\varepsilon)$ in $D_{r'}$ for some $0 < r' \leqq r$?

Consider the following example:

$$A(\varepsilon) = \begin{pmatrix} 1 & \varepsilon \\ 0 & 1 \end{pmatrix}, \qquad B(\varepsilon) = \begin{pmatrix} 1 & \varepsilon^2 \\ 0 & 1 \end{pmatrix}. \tag{1.5}$$

Clearly $A(\varepsilon) \sim_p B(\varepsilon)$ in **C**. On the other hand $A(\varepsilon) \not\sim_a B(\varepsilon)$ in any $D_{r'}$ $(r' > 0)$. Otherwise

$$\begin{aligned} A_1(\varepsilon) &= \varepsilon^{-1}(A(\varepsilon) - I) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \sim_a, \\ B_1(\varepsilon) &= \varepsilon^{-1}(B(\varepsilon) - I) = \begin{pmatrix} 0 & \varepsilon \\ 0 & 0 \end{pmatrix}, \qquad |z| < r'. \end{aligned} \tag{1.6}$$

But this is impossible since $A_1(0)$ and $B_1(0)$ are not similar. This shows that the above problem does not have a simple solution.

Wasow [3] gave a simple condition when pointwise similarity implies analytic similarity in the neighborhood of the origin. Consider the matrix equation

$$A(\varepsilon)X - XA(\varepsilon) = 0. \tag{1.7}$$

Of course, we can view (1.7) as a system of n^2 linear homogeneous equations in n^2 unknowns x_{ij}, $i, j = 1, \cdots, n$ $(X = (x_{ij})_1^n)$. Fix ε and let $\kappa(\varepsilon)$ be the number of linearly independent solutions of (1.7). $\kappa(\varepsilon)$ can be easily determine by the degrees of the invariant polynomials of $A(\varepsilon)$ (e.g. [1, ch. 8, sec. 2]). It is not difficult to see that there exists $0 < \rho$ such that

$$\kappa(\varepsilon) = \kappa, \qquad 0 < |\varepsilon| < \rho. \tag{1.8}$$

WASOW'S CONDITION [3]. *Assume that*

$$\kappa(0) = \kappa. \tag{1.9}$$

Then $A(\varepsilon) \sim_a B(\varepsilon)$ *in* $D_{r'}$ *if and only if* $A(\varepsilon) \sim_p B(\varepsilon)$ *in* D_r.

The aim of this paper is to give conditions under which the pointwise similarity implies holomorphic similarity in case that Wasow's condition fails. The starting point of our investigation is the following theorem.

THEOREM 2.1. *Let* $A(\varepsilon)$ *and* $B(\varepsilon)$ *be* $n \times n$ *matrices analytic in* ε *for* $|\varepsilon| < r$. *There exists a non-negative integer* ω *depending only on* $A(\varepsilon)$ *such that* $A(\varepsilon) \sim_a B(\varepsilon)$ *for* $\varepsilon \in D_{r'}$ $(r' > 0)$ *if and only if* $A(\varepsilon) \sim_p B(\varepsilon)$ *for* $\varepsilon \in D_{r'}$ *and there exists* $R(\varepsilon)$ *of the form*

$$R(\varepsilon) = \sum_{k=0}^{\omega} R_k \varepsilon^k, \qquad R_k \in M_n(\mathbf{C}), \quad |R_0| \neq 0 \tag{1.10}$$

such that

$$A(\varepsilon)R(\varepsilon) - R(\varepsilon)B(\varepsilon) = \varepsilon^{\omega+1}O(1). \tag{1.11}$$

We determine an explicit upper bound for ω. We also give a simple sufficient criterion which implies that the conditions (1.10) and (1.11) for $\omega = 1$ guarantee a positive answer to our problem. In Section 3 we examine the conditions (1.10)–(1.11) for $\omega = 1$. This problem leads us to the notion of conjugacy of two matrices X and Y with respect to a matrix Z. In case that $Z = cI$ this is the standard notion of similarity. We give a procedure to determine when X and Y are conjugate with respect to Z and in some cases the verification is quite straightforward. However, the solution of the general problem is incomplete. In Section 4 we show how to determine whether (1.11) is solvable. In fact (1.10)–(1.11) is equivalent to the notion of strong similarity of certain upper block triangular matrices. We also give a simple necessary and sufficient condition for the solution of Problem 1.1 for certain type of matrices $A(\varepsilon)$ which do not satisfy the Wasow condition.

THEOREM 4.2. *Let* $A(\varepsilon)$ *be complex valued matrix analytic in* ε *at the origin. Assume that the Wasow condition fails. Suppose that the subspace of all matrices* R_0 *which satisfy*

$$R_0A_0 = A_0R_0, \qquad \operatorname{tr}[V(R_0A_1 - A_1R_0] = 0 \tag{1.12}$$

for all V *which commute with* A_0, *is of dimension* κ. *Then* $A(\varepsilon) \sim_a B(\varepsilon)$ *if and only if there exists a nonsingular matrix* P *commuting with* A_0 *and a matrix* R *such that*

$$PB_1 - A_1P = A_0R - RA_0 \tag{1.13}$$

provided that $B(\varepsilon)$ is normalized by the condition

$$B_0 = A_0. \tag{1.14}$$

That is A_1 and B_1 are conjugate with respect to A_0.

We state a conjecture which determines the smallest ω described on Theorem 2.1. In fact Theorem 4.2 supports this conjecture. In the last section we show that (1.10)-(1.11) for any ω is equivalent to the same problem with $\omega = 1$ stated for appropriate choice of matrices A_0', A_1' and B_0', B_1'.

2. Main results

PROOF OF THEOREM 2.1. Assume that the Wasow condition holds. Then the pointwise similarity implies analytic similarity. In that case the value of the ω is zero. Indeed, as $A(0) \sim B(0)$ there exists non-singular R_0 such that

$$B(0) = R_0^{-1} A(0) R_0. \tag{2.1}$$

Now

$$A(\varepsilon)R_0 - R_0 B(\varepsilon) = \varepsilon O(1). \tag{2.2}$$

as we claimed.

Suppose now that the Wasow condition fails. That is

$$\kappa < \kappa(0). \tag{2.3}$$

Rewrite the system (1.7) as a system of linear equations in n^2 unknowns x_{ij}, $i, j = 1, \cdots, n$,

$$\hat{A}(\varepsilon)\hat{X} = 0. \tag{2.4}$$

Here $\hat{A}(\varepsilon)$ is an $n^2 \times n^2$ matrix

$$\begin{gathered} \hat{A}(\varepsilon) = (\hat{a}_{(i,j),(p,q)}(\varepsilon)), \qquad \hat{X} = (x_{(i,j)}) - (\text{vector}), \\ \hat{a}_{(i,j),(p,q)} = a_{ip}\delta_{qj} - \delta_{ip}a_{qj}, \qquad i, j, p, q = 1, \cdots, n. \end{gathered} \tag{2.5}$$

Using the tensor product one can write

$$\hat{A}(\varepsilon) = I \otimes A(\varepsilon) - A^t(\varepsilon) \otimes I, \qquad (A^t \text{ — the transpose of } A). \tag{2.6}$$

See, for example [2, p. 8]. The condition (2.3) implies the existence of an $\eta \times \eta$ submatrix of $\hat{A}(\varepsilon)$ — call it $P(\varepsilon)$ such that

$$|P(\varepsilon)| = a\varepsilon^s(1 + \varepsilon O(1)), \qquad a \neq 0. \tag{2.7}$$

Here by $|P|$ we denote the determinant of a square matrix and

$$\eta = n^2 - \kappa. \tag{2.8}$$

We claim that if one can satisfy the conditions (1.10) and (1.11) with $\omega = s$ then $A(\varepsilon) \sim_a B(\varepsilon)$. Indeed, assume that $\omega = s$ and (1.10) and (1.11) holds. Since $|R_0| \neq 0$ there exists $0 < r' \leqq r$ such that $R(\varepsilon)^{-1}$ exists for $|\varepsilon| < r'$. Let

$$C(\varepsilon) = R(\varepsilon)B(\varepsilon)R(\varepsilon)^{-1}. \tag{2.9}$$

Clearly it is enough to show that $A(\varepsilon) \sim_a C(\varepsilon)$. Also $A(\varepsilon) \sim_p C(\varepsilon)$. Consider the system

$$A(\varepsilon)Y - YC(\varepsilon) = 0. \tag{2.10}$$

Rewrite (2.10) in the form of the system in n^2 variables

$$F(\varepsilon)\hat{Y} = 0. \tag{2.11}$$

In tensor notation

$$F(\varepsilon) = I \otimes A(\varepsilon) - C'(\varepsilon) \otimes I \tag{2.12}$$

According to our assumptions

$$A(\varepsilon) - C(\varepsilon) = \varepsilon^{s+1}O(1). \tag{2.13}$$

So

$$F(\varepsilon) - \hat{A}(\varepsilon) = (A'(\varepsilon) - C'(\varepsilon)) \otimes I = \varepsilon^{s+1}O(1). \tag{2.14}$$

Consider the submatrix $P(\varepsilon)$ of $\hat{A}(\varepsilon)$. Assume that the η rows of $P(\varepsilon)$ form the set $J \subset \mathcal{N} \times \mathcal{N}$ ($\mathcal{N} = \{1, 2, \cdots, n\}$) and the η columns of $P(\varepsilon)$ form the set $K \subset \mathcal{N} \times \mathcal{N}$. Look at the corresponding submatix $Q(\varepsilon)$ of $F(\varepsilon)$ which is formed by the rows J and the columns K. From (2.14) and (2.7) it follows that

$$|Q(\varepsilon)| = a\varepsilon^s(1 + \varepsilon O(1)). \tag{2.15}$$

As $C(\varepsilon)$ is pointwise similar to $A(\varepsilon)$ we must have that the system (2.10) has the same number of linearly independent solutions as (1.7). Therefore any $(\eta + 1) \times (\eta + 1)$ minor of $F(\varepsilon)$ vanishes. Let $Y(\varepsilon)$ be the unique solution of (2.10) satisfying the conditions

$$y_{ij}(\varepsilon) = \delta_{ij} \qquad \text{if } (i, j) \notin K. \tag{2.16}$$

We assert that $Y(\varepsilon)$ is holomorphic at $\varepsilon = 0$ and

$$Y(0) = I. \tag{2.17}$$

Indeed consider the unique solution $X(\varepsilon)$ of (1.7) satisfying the condition (2.16). Clearly $X(\varepsilon) = I$ is this solution. Using the Cramer formulas for the solutions of (1.7) and (2.10) (only to the equations corresponding to the entries (i, j), $(i, j) \in J$) and taking in account (2.7), (2.15) and (2.14) we get

$$Y(\varepsilon) = (1 + \varepsilon O(1))X(\varepsilon) + \varepsilon O(1). \tag{2.18}$$

This establishes (2.17) and the analyticity of Y around the neighborhood of the origin. So there exists $0 < r'' \leqq r'$ such that $Y(\varepsilon)$ and $Y(\varepsilon)^{-1}$ are holomorphic in $|\varepsilon| < r''$. This proves the existence of ω depending only on $A(\varepsilon)$ such that (1.10), (1.11) together with the assumptions $A(\varepsilon) \sim_p B(\varepsilon)$ at the origin imply that $A(\varepsilon) \sim_a B(\varepsilon)$ at $\varepsilon = 0$. Vice versa, if $A(\varepsilon) \sim_a B(\varepsilon)$ for $\varepsilon \in D_{r'}$ then $A(\varepsilon) \sim_p B(\varepsilon)$ for $\varepsilon \in D_{r'}$ and (1.10) and (1.11) hold for any integer. The proof of the theorem is completed.

DEFINITION 2.1. Let $A(\varepsilon)$ be complex valued matrix analytic in ε at the origin. Then μ is called the minimal index of $A(\varepsilon)$ at $\varepsilon = 0$ if Theorem 2.1 holds for $\omega = \mu$, but if $\omega < \mu$ then there exists $B(\varepsilon)$ which satisfies the conditions of Theorem 2.1 but (1.10) and (1.11) do not imply that $A(\varepsilon) \sim_a B(\varepsilon)$.

As we pointed out in the proof of Theorem 2.1 Wasow's condition (1.9) implies that $\mu = 0$. From the proof of Theorem 2.1 we deduce

THEOREM 2.2. *Let η be given by* (2.8) *and consider all non-zero $\eta \times \eta$ minors of $I \otimes A(\varepsilon) - A'(\varepsilon) \otimes I$ which must be of the form* (2.7). *Let ν be the minimum of all possible exponents s appearing in* (2.7). *Then*

$$\mu \leqq \nu. \tag{2.19}$$

Clearly $\nu = 0$ if and only if the Wasow condition (1.9) holds. Next we give a sufficient condition for $\mu = 1$

THEOREM 2.3. *Let $A(\varepsilon)$ satisfy the assumptions of Theorem* 2.1. *Assume that the Wasow condition fails* (*i.e.* (2.3) *holds*). *Suppose that ν given in Theorem* 2.2 *equals*

$$\nu = \kappa(0) - \kappa. \tag{2.20}$$

Then the minimal index of $A(\varepsilon)$ at the origin does not exceed 1.

To prove this theorem we need the following lemma.

LEMMA 2.1. *Let X be an $n \times n$ matrix whose rank is k $(\leqq n)$. Then for any $n \times n$ matrix Y and analytic valued $n \times n$ matrix $Z(\varepsilon)$ $(|\varepsilon| < r)$ the following relations hold*:

$$|X + \varepsilon Y| = \varepsilon^{n-k}O(1), \tag{2.21}$$

$$|X + \varepsilon Y + \varepsilon^2 Z(\varepsilon)| = |X + \varepsilon Y| + \varepsilon^{n-k+1}O(1). \tag{2.22}$$

PROOF. Let $A(\varepsilon) = (a_{ij}(\varepsilon))_1^n$ be an analytic valued matrix at $\varepsilon = 0$. Let $r = (r_1, \cdots, r_n)$ be a vector with non-negative integer coordinates. As usual denote $|r| = \sum_{i=1}^n r_i$. By $(a_{ij}^{(r_i)}(\varepsilon))_1^n$ denote the matrix whose i-th row is the r_i-th derivative of $A(\varepsilon)$. From the standard formula of the derivative of the determinant we deduce

$$\frac{d^p}{d\varepsilon^p}|A(\varepsilon)| = \sum_{|r|=p} \frac{p!}{r_1! \cdots r_n!} |(a_{ij}^{(r_i)}(\varepsilon))_1^n|. \tag{2.23}$$

Put

$$A(\varepsilon) = X + \varepsilon Y + \varepsilon^2 Z(\varepsilon) \tag{2.24}$$

and let $\varepsilon = 0$ in (2.23). Set

$$G = (a_{ij}^{(r_i)}(0))_1^n, \quad \sum_{i=1}^n r_i = p, \quad r_{i_1} = \cdots = r_{i_q} = 0, \quad r_j > 0 \text{ if } j \neq i_1, \cdots, i_q. \tag{2.25}$$

Let

$$G\begin{pmatrix} i_1, \cdots, i_q \\ j_1, \cdots, j_q \end{pmatrix}$$

be a $q \times q$ minor of G composed of $i_1, \cdots, i_q$ rows and $j_1, \cdots, j_q$ columns of G. In view of (2.25) we have

$$G\begin{pmatrix} i_1, \cdots, i_q \\ j_1, \cdots, j_q \end{pmatrix} = X\begin{pmatrix} i_1, \cdots, i_q \\ j_1, \cdots, j_q \end{pmatrix}, \qquad 1 \leqq \begin{matrix} i_1 \\ j_1 \end{matrix}, \cdots, \begin{matrix} i_k \\ j_k \end{matrix} \leqq n. \tag{2.26}$$

Assume first that $p < n - k$. Then $q \geqq k + 1$ and since $r(X) = k$ both sides of (2.26) are equal to zero. Expanding the determinant of G by the rows $i_1, \cdots, i_q$ we obtain that $|G| = 0$. So

$$\frac{d^p}{d\varepsilon^p}|A(\varepsilon)|\bigg|_{\varepsilon=0} = 0, \qquad p = 0, \cdots, n-k-1. \tag{2.27}$$

Assume now that $p = n - k$. Again if $q \geqq k + 1$, $|G| = 0$. So we are left with the case where $q = k$. That is, there exist $1 \leqq i_1' < i_2' < \cdots < i_{n-k}' \leqq n$ such that

$$r_{i_1'} = \cdots = r_{i_{n-k}'} = 1. \tag{2.28}$$

In this case G is composed of $i_1', \cdots, i_{n-k}'$ rows of Y and $i_1, \cdots, i_k$ rows of X. Therefore we showed

$$\frac{d^{n-k}}{d\varepsilon^{n-k}} |A(\varepsilon)| \Big|_{\varepsilon=0} = \frac{d^{n-k}}{d\varepsilon^{n-k}} |X + \varepsilon Y| \Big|_{\varepsilon=0}. \tag{2.29}$$

This verifies (2.21) and (2.22).

PROOF OF THEOREM 2.3. Assume that $B(\varepsilon) \sim_p A(\varepsilon)$ for $\varepsilon \in D_r$. Suppose that (1.10) and (1.11) hold for $\omega = 1$. We claim that $B(\varepsilon) \sim_a A(\varepsilon)$ for $\varepsilon \in D_{r'}$ if (2.20) holds. Our proof is a modified version of the proof of Theorem 2.1. We just point out the arguments which should be modified. According to (2.20) and the definition of ν we may assume that s given in (2.7) equals ν. From (2.6), (2.12) and the equality $\omega = 1$ we get

$$\begin{aligned} \hat{A}(\varepsilon) &= (I \otimes A_0 - A_0^t \otimes I) + \varepsilon(I \otimes A_1 - A_1^t \otimes I) + \varepsilon^2 O(1), \\ F(\varepsilon) &= (I \otimes A_0 - A_0^t \otimes I) + \varepsilon(I \otimes A_1 - A_1^t \otimes I) + \varepsilon^2 O(1). \end{aligned} \tag{2.30}$$

Thus we can apply Lemma 2.1 to the $J \times K$ minors of $\hat{A}(\varepsilon)$ and $F(\varepsilon)$. So

$$|Q(\varepsilon)| - |P(\varepsilon)| = \varepsilon^{\eta - \eta(0) + 1} O(1), \qquad \eta(0) = \eta^2 - \kappa(0). \tag{2.31}$$

This establishes (2.15). It is left to show (2.18). Use again the Cramer formulas for the solutions of (1.7) and (2.10) (only for the equations corresponding to the entries (i, j), $(i, j) \in J$). Thus we have to consider $\eta \times \eta$ minors consisting of $\eta - 1$ columns of $\hat{A}(\varepsilon)$ $(F(\varepsilon))$ from the set K and a column which is a linear combination of the columns of $\hat{A}(\varepsilon)$ $(F(\varepsilon))$ which do not belong to K. Clearly the rank of such a minor at $\varepsilon = 0$ is at most $\eta(0)$. Using (2.31) and (2.22) we obtain that the difference between the corresponding minors of $\hat{A}(\varepsilon)$ and $F(\varepsilon)$ is at least of the form $\varepsilon^{\eta - \eta(0) + 1} O(1)$, i.e. $\varepsilon^{\nu + 1} O(1)$. Dividing the minors of $\hat{A}(\varepsilon)$ by $|P(\varepsilon)|$ and the minors of $F(\varepsilon)$ by $|Q(\varepsilon)|$ from (2.7) and (2.15) we deduce (2.18). The proof of the theorem is completed.

3. The case $\omega = 1$

Assume that $A(\varepsilon)$ and $B(\varepsilon)$ are analytic valued at the origin and have the expansions (1.1). Assume that $A(\varepsilon) \sim_a B(\varepsilon)$ for $\varepsilon \in D_r$. In particular $A(0)$ is similar to $B(0)$. By considering $TB(\varepsilon)T^{-1}$ for a suitable $T \in M_n(\mathrm{C})$ we may assume in (1.1) that

$$A_0 = B_0. \tag{3.1}$$

In that case the conditions (1.10) and (1.11) for $\omega = 1$ are equivalent to

$$A_0R_0 - R_0A_0 = 0, \ |R_0| \neq 0, \tag{3.2}$$

$$A_0R_1 + A_1R_0 - R_1A_0 - R_0B_1 = 0. \tag{3.3}$$

Definition 3.1. Let $X, Y, Z \in M_n(\mathbf{C})$. The matrix X is conjugate to Y with respect to Z, if there exists a non-singular matrix P commuting with Z

$$ZP - PZ = 0, \tag{3.4}$$

such that

$$XP - PY = ZQ - QZ \tag{3.5}$$

for some $Q \in M_n(\mathbf{C})$.

Denote this relation by $X \sim Y(Z)$. Clearly, if $Z = cI$ then X is conjugate to Y if and only if X is similar to Y. It is easy to check that for a fixed Z the relation $X \sim Y(Z)$ is an equivalence relation. Thus, the problem of determining whether (3.2)–(3.3) are solvable is equivalent to the problem whether $A_1 \sim B_1(A_0)$. In this section we shall give a partial answer to the following problem.

Problem 3.1. Given $X, Y, Z \in M_n(\mathbf{C})$, find necessary and sufficient conditions for X to be conjugate to Y with respect to Z.

Clearly this problem makes sense if $X, Y, Z \in M_n(\mathcal{F})$ for any field $\mathcal{F}$. We shall restrict ourselves to the field of complex numbers although our approach will apply for any field $\mathcal{F}$. Our first observation is

Lemma 3.1. *Let $U, Z \in M_n(\mathbf{C})$. Then U is a commutator of Z and Q, i.e.*

$$U = ZQ - QZ \tag{3.6}$$

for some Q, if and only if

$$\operatorname{tr}(VU) = 0 \tag{3.7}$$

for any V which commutes with Z. (Here $\operatorname{tr}(W)$ denotes the trace of W.)

Proof. Clearly if V commutes with Z then

$$\begin{aligned} \operatorname{tr}(VU) &= \operatorname{tr}(VZQ - VQZ) = \operatorname{tr}[Z(VQ) - VQZ] \\ &= \operatorname{tr}[(VQ)Z - VQZ] = 0. \end{aligned} \tag{3.8}$$

Vice versa, suppose that (3.7) holds for any V which commutes with Z. Consider the equality (3.6) as a system of n^2 non-homogeneous equations in the unknowns q_{ij}, $i, j = 1, \cdots, n$ $(Q = (q_{ij})_1^n)$. In tensor form (3.6) is given as

$$(I \otimes Z - Z^t \otimes I)\hat{Q} = \hat{U} \tag{3.9}$$

if we adopt the notation of the previous section. It is well known that (3.9) is solvable if and only if $\hat{U}$ is orthogonal to any solution of the adjoint system. That is

$$0 = \sum_{i,j=1}^{n} w_{ij}u_{ij} = \operatorname{tr}(W^t U), \quad W = (w_{ij})_1^n, \quad U = (u_{ij})_1^n, \tag{3.10}$$

$$(I \otimes Z - Z^t \otimes I)^t \hat{W} = (I \otimes Z^t - Z \otimes I)\hat{W} = 0. \tag{3.11}$$

Now (3.11) means that

$$Z^t W - W Z^t = 0. \tag{3.12}$$

Thus W^t commutes with Z and (3.10) is equivalent to (3.7). End of proof.

Let $V_1, \cdots, V_k$ form a basis for the subspace of all matrices in $M_n(\mathbf{C})$ which commute with Z. So any P which satisfies (3.4) is of the form

$$P = \sum_{i=1}^{k} v_i V_i. \tag{3.13}$$

According to Lemma 3.1 (3.5) is solvable for some Q if and only if

$$\operatorname{tr}[V_j(XP - PY)] = 0, \qquad j = 1, \cdots, k. \tag{3.14}$$

The equations (3.13)–(3.14) determine the subspace $\mathcal{P}$ of all matrices P which solve (3.4)–(3.5). It is left to find whether $\mathcal{P}$ contains a non-singular matrix.

In principle this can be done by verifying a finite number of conditions. Indeed let $\mathcal{P}$ be any subspace of matrices P of the form (3.13). Define

$$F(v_1, \cdots, v_k) = \left| \sum_{i=1}^{k} v_i V_i \right| = \sum_{|p|=n} a_p v^p, \quad p = (p_1, \cdots, p_k), \quad v^p = v_1^{p_1} \cdots v_k^{p_k}. \tag{3.15}$$

Thus $\mathcal{P}$ does not contain a non-singular matrix if and only if F is zero identically. It is a standard fact that a polynomial F of degree n is zero identically if and only if F vanishes at the test points

$$v_i = 0, 1, \cdots, n, \qquad i = 1, \cdots, k. \tag{3.16}$$

Moreover the number of test points can be reduced by observing that

$$F(tv_1, \cdots, tv_k) = t^n F(v_1, \cdots, v_k). \tag{3.17}$$

Next we observe that

$$X \sim Y(Z) \quad \text{if and only if} \quad TXT^{-1} \sim TYT^{-1}(TZT^{-1}). \tag{3.18}$$

Since we are working over $M_n(\mathbf{C})$ we may assume that Z is in the Jordan canonical form

$$Z = \operatorname{diag}\{J_1, \cdots, J_u\}, \quad J_k = \lambda_k I_k + H_k, \quad \dim J_k = n_k, \quad k = 1, \cdots, u. \tag{3.19}$$

Here I_k is the identity matrix and H_k the 0–1 matrix whose non-zero elements are on the upper diagonal. In that case the subspace of all commuting matrices P with Z is well known (e.g. [1, ch. 8, sec. 1]).

LEMMA 3.2. *Let $Z \in M_n(\mathbf{C})$ be a matrix given by* (3.19). *Then a block matrix $P = (P_{\alpha\beta})_1^u \in M_n(\mathbf{C})$ commutes with Z if and only if the blocks $P_{\alpha\beta} = (p_{ij}^{(\alpha\beta)})$, $i = 1, \cdots, n_\alpha$, $j = 1, \cdots, n_\beta$ satisfy the following conditions*:

$$\text{if } \lambda_\alpha \neq \lambda_\beta \quad p_{\alpha\beta} = 0, \tag{3.20}$$

$$\text{if } \lambda_\alpha = \lambda_\beta \quad p_{ij}^{(\alpha\beta)} = 0 \quad \text{for } j < i + n_\beta - \min(n_\alpha, n_\beta), \tag{3.21}$$

$$p_{ij}^{(\alpha\beta)} = p_{(i+1)(j+1)}^{(\alpha\beta)} \quad \text{for } j \geqq i + n_\beta - \min(n_\alpha, n_\beta).$$

In fact if $n_1, \cdots, n_t$ are the degrees of the non-constant invariant polynomials $i_1(\lambda), \cdots, i_t(\lambda)$ of Z then the number of free parameters in P is

$$N = \sum_{i=1}^{t} (2i-1) n_i. \tag{3.22}$$

Applying Lemmas 3.1 and 3.2 we obtain

LEMMA 3.3. *Assume that Z is of the form* (3.19). *Then P solves* (3.4) *and* (3.5) *if and only if for any two indices α, β such that $\lambda_\alpha = \lambda_\beta$ and any $V_{\alpha\beta}$ of the form* (3.21) *the following equality holds*:

$$0 = \operatorname{tr}\left\{ V_{\alpha\beta} \left[\sum_{j=1}^{u} (X_{\beta j} P_{j\alpha} - P_{\beta j} Y_{j\alpha}) \right] \right\}, \tag{3.23}$$
$$X = (X_{ij})_1^u, \qquad Y = (Y_{ij})_1^u$$

provided that P is of the form (3.20)–(3.21).

Noting that $V_{j\alpha} = P_{j\alpha} = 0$ if $\lambda_j \neq \lambda_\alpha$ we deduce

THEOREM 3.1. *Assume that Z is of the form*

(3.24) $$Z = \operatorname{diag}\{Z_1, \cdots, Z_v\}$$

such that

(3.25) $$(\mu_j I_j - Z_j)^n = 0, \qquad \mu_j \neq \mu_k, \quad j \neq k, \quad j, k = 1, \cdots, v.$$

Then X is conjugate to Y with respect to Z if and only if

(3.26) $$X_{ii} \sim Y_{ii}(Z_i), \quad i = 1, \cdots, v, \quad X = (X_{ij})_1^v, \quad Y = (Y_{ij})_1^v.$$

Thus in Problem 3.1 we may assume that Z is a nilpotent matrix. In case that Z is similar to a diagonal matrix then Problem 3.1 has a simple solution.

COROLLARY 3.1. *Let the assumptions of Theorem 3.1 hold. Assume furthermore that Z is similar to a diagonal matrix. Then $X \sim Y(Z)$ if and only if X_{ii} is similar to Y_{ii} for $i = 1, \cdots, v$.*

THEOREM 3.2. *Assume that Z consists of one Jordan block*

(3.27) $$Z = H, \quad H = (h_{ij})_1^n, \quad h_{ij} = \delta_{(i+1)j}, \qquad i, j = 1, \cdots, n.$$

Then $X = (x_{ij})_1^n$ is similar to $Y = (y_{ij})_1^n$ with respect to Z if and only if

(3.28) $$\sum_{k=1}^{n-i} x_{(i+k)k} = \sum_{k=1}^{n-1} y_{(i+k)k}, \qquad i = 0, \cdots, n-1.$$

PROOF. It is a well known fact that any P which commutes with Z given by (3.27) is a polynomial in H:

(3.29) $$P = \sum_{i=0}^{n-1} a_i H^i.$$

The assumption that P is nonsingular is equivalent to the fact that $a_0 \neq 0$. So we may assume that $a_0 = 1$. Then the condition (3.14) states

(3.30) $$0 = \operatorname{tr}(H^j XP - H^j PY) = \operatorname{tr}(XPH^j - YH^j P) = \operatorname{tr}\left[(X - Y)\left(\sum_{i=0}^{n-j-1} a_i H^{i+j}\right)\right], \qquad j = n-1, \cdots, 0.$$

For $j = n - 1$ (3.30) is equivalent to

(3.31) $$\operatorname{tr}[(X - Y)H^j] = 0.$$

Assume that we already proved (3.31) for $j = n - 1, \cdots, k$. By letting j in (3.30) be $k - 1$ we deduce that X and Y satisfy (3.31) for $k - 1$. So (3.31) holds for $j = n - 1, \cdots, 0$. This is exactly the conditions (3.28). Conversely if (3.28) hold then (3.30) is fulfilled when $P = I$. So $X \sim Y(Z)$. The proof of the theorem is completed.

4. The general problem

The conditions (1.11) can be stated in terms of matrix equalities

$$A_0R_k - R_kA_0 = \sum_{i=1}^{k} (R_{k-i}B_i - A_iR_{k-i}), \tag{4.1}$$

for $k = 0, 1, \cdots, \omega$, where we assumed the normalization $A_0 = B_0$. A sequence $(R_0, \cdots, R_j)$ is called a solution if $R_0, \cdots, R_j$ satisfy (4.1) for $k = 0, 1, \cdots, j$. Denote by $\mathscr{L}_j$ the subspace of all solutions $(R_0, \cdots, R_j)$ and by $\mathscr{L}_{j,i}$ the subspace of the first i matrices $(R_0, \cdots, R_i)$ in the solutions $(R_0, \cdots, R_j)$ where $0 \leqq i \leqq j$. Clearly

$$\mathscr{L}_{j,i} \supseteq \mathscr{L}_{j+1,i}. \tag{4.2}$$

According to Lemma 3.1 $\mathscr{L}_{j+1,j}$ is the subspace of all solutions $(R_0, \cdots, R_j)$ such that

$$\operatorname{tr}\left[V\sum_{i=1}^{j+1} (R_{j+1-i}B_i - A_iR_{j+1-i})\right] = 0, \qquad VA_0 = A_0V, \tag{4.3}$$

for all V which commute with A_0. Thus if we constructed $\mathscr{L}_j$ (4.3) determines $\mathscr{L}_{j+1,j}$. Now by solving (4.1) for $k = j+1$ where $(R_0, \cdots, R_j) \in \mathscr{L}_{j+1,j}$ we obtain the subspace $\mathscr{L}_{j+1}$. Thus if $A(\varepsilon) \sim_p B(\varepsilon)$ then $A(\varepsilon) \sim_a B(\varepsilon)$ if and only if $\mathscr{L}_{\nu,0}$ contains a non-singular matrix (ν is given in Theorem 2.2).

THEOREM 4.1. *Assume that $A(\varepsilon)$ and $B(\varepsilon)$ are analytically similar at the origin. Consider the system* (4.1) *for $k = 0, \cdots, j$. Then*

$$\dim \mathscr{L}_{j,0} \geqq \kappa \tag{4.4}$$

for any $j \geqq 0$. Moreover the equality sign holds if j is not less ν (given in Theorem 2.2).

To prove this theorem we need the following lemma.

LEMMA 4.1. *Let $A(\varepsilon)$ be a complex valued matrix analytic in ε at the origin. Consider all complex valued matrices $X(\varepsilon) = \sum_{k=0}^{\infty} X_k\varepsilon^k$ analytic in ε at the origin and satisfying the equation* (1.7). *Then the set of all possible X_0 form a subspace U of dimension κ.*

PROOF. First we claim

$$\dim U \leqq \kappa.$$

Indeed let $X^{(1)}(\varepsilon), \cdots, X^{(\kappa+1)}(\varepsilon)$ be $\kappa+1$ analytic solutions of (1.7). Let $G(\varepsilon)$ be $n^2 \times (\kappa+1)$ matrix whose columns are the vectors $X^{(1)}(\varepsilon), \cdots, X^{(\kappa+1)}(\varepsilon)$. By the definition of κ, $X^{(1)}(\varepsilon), \cdots, X^{(\kappa+1)}(\varepsilon)$ are linearly dependent. So $r(G(\varepsilon))$ — the rank of $G(\varepsilon)$ — satisfies $r(G(\varepsilon)) \leqq \kappa$. In particular $r(G(0)) \leqq \kappa$ which proves the assertion. Next we show the existence of κ analytic solutions $X^{(1)}(\varepsilon), \cdots, X^{(\kappa)}(\varepsilon)$ of (1.7) which are linearly independent for $0 < |\varepsilon|$. We follow the notation in the proof of Theorem 2.1. So all $(\eta+1) \times (\eta+1)$ $(\eta = n^2 - \kappa)$ minors of $\hat{A}(\varepsilon)$ (2.6) vanish identically and there exist $\eta \times \eta$ minors $P(\varepsilon)$ of the form (2.7).

Let K' be the complementary set of K in $\mathcal{N} \times \mathcal{N}$. For $\alpha \in K'$ define $Y^{(\alpha)}(\varepsilon) = (y_{ij}^{(\alpha)}(\varepsilon))_1^n$ to be the following unique solution of (1.7):

$$y_{ij}^{(\alpha)}(\varepsilon) = \varepsilon^s \quad \text{if } (i,j) = \alpha, \qquad y_{ij}^{(\alpha)}(\varepsilon) = 0 \quad \text{if } \alpha \neq (i,j) \in K'. \tag{4.5}$$

From the proof of Theorem 2.1 it follows that $Y^{(\alpha)}(\varepsilon)$ are analytic. Clearly $(Y^{(\alpha)}(\varepsilon))$, $\alpha \in K'$, are linearly independent for $|\varepsilon| > 0$. Let $H(\varepsilon)$ be an $n^2 \times \kappa$ matrix whose columns are vectors $X^{(1)}(\varepsilon), \cdots, X^{(\kappa)}(\varepsilon)$ which are analytic solutions of (1.7). Assume that $r(H(\varepsilon)) = \kappa$. If $r(H(0)) = \kappa$ we have finished the proof. Assume that $r(H(0)) < \kappa$. So there exists $\kappa \times \kappa$ minor of $H(\varepsilon)$ of the form

$$|Q(\varepsilon)| = a'\varepsilon^{s'}(1 + \varepsilon O(1)), \qquad a' \neq 0, \quad s' \geqq 1. \tag{4.6}$$

As $X^{(1)}(0), \cdots, X^{(\kappa)}(0)$ are linearly dependent we have

$$\sum_{i=1}^{\kappa} \alpha_i X^{(i)}(0) = 0. \tag{4.7}$$

For simplicity of notation we may assume that $\alpha_\kappa = 1$. Consider a new set $\bar{X}^{(1)}(\varepsilon), \cdots, \bar{X}^{(\kappa)}(\varepsilon)$ of linearly independent analytic solutions of (1.7):

$$\bar{X}^{(i)}(\varepsilon) = X^{(i)}(\varepsilon), \quad i = 1, \cdots, \kappa - 1, \qquad \bar{X}^{(\kappa)} = \varepsilon^{-1} \sum_{i=1}^{\kappa} \alpha_i X^{(i)}(\varepsilon). \tag{4.8}$$

Let $\bar{H}(\varepsilon)$ be the matrix composed of $\bar{X}^{(1)}, \cdots, \bar{X}^{(\kappa)}$. Again if $r(\bar{H}(0)) = \kappa$ we are done. Otherwise considering the corresponding minor $\bar{Q}(\varepsilon)$ which consists of the same rows and columns as $Q(\varepsilon)$, we easily deduce that

$$|\bar{Q}(\varepsilon)| = a'\varepsilon^{s'-1}(1 + \varepsilon O(1)). \tag{4.9}$$

Continuing in the same manner we shall finally deduce the lemma.

PROOF OF THEOREM 4.1. Let $X(\varepsilon)$ be an analytic solution of (1.7). Denote

(4.10) $$R(\varepsilon) = X(\varepsilon)T^{-1}(\varepsilon)$$

where $T(\varepsilon)$ satisfies (1.4). Thus

(4.11) $$A(\varepsilon)R(\varepsilon) - R(\varepsilon)B(\varepsilon) = 0.$$

We also have

(4.12) $$R_0 = X_0 T_0^{-1}.$$

As $R(\varepsilon) = \Sigma_{k=0}^{\infty} R_k \varepsilon^k$ (4.1) is satisfied for $k = 0, 1, 2, \cdots$. From Lemma 4.1 we deduce the inequality (4.4). To finish the proof of the theorem we have to verify the equality

(4.13) $$\dim \mathcal{L}_{\nu,0} = \kappa.$$

Assume that $R(\varepsilon) = \Sigma_{k=0}^{\infty} R_k \varepsilon^k$ satisfy (1.11). Here we do not demand that $|R_0| \neq 0$. Moreover assume that $\omega = \nu$. Define $X(\varepsilon)$ by the equation (4.10). From (1.11) and (1.4) we get

$$A(\varepsilon)X(\varepsilon) - X(\varepsilon)A(\varepsilon) = \varepsilon^{\nu+1}O(1).$$

Repeating the arguments of the proof of Theorem 2.1 we obtain the existence of the unique analytic solution $Y(\varepsilon)$ of (1.7) such that $x_{ij}(\varepsilon) = y_{ij}(\varepsilon)$ if $(i, j) \in K'$. Moreover $X(0) = Y(0)$. This manifests that $\dim U \geqq \dim \mathcal{L}_{\nu,0} \geqq \kappa$. Now Lemma 4.1 implies (4.13). The proof of theorem is completed.

Theorem 4.1 can be obviously applied to the case $B(\varepsilon) = A(\varepsilon)$.

Definition 4.1. Consider the system of matrix equations

(4.14) $$A_0 R_k - R_k A_0 = \sum_{i=1}^{k} R_{k-i} A_i - A_i R_{k-i}$$

for $k = 0, 1, \cdots, j$. Let U_j be the subspace spanned by the matrices R_0 in the solutions $(R_0, \cdots, R_j)$. Define μ' to be the following non-negative integer:

(4.15) $$U_{\mu'-1} \supsetneqq U, \qquad U_{\mu'} = U$$

where U is given by Lemma 4.1.

Theorem 4.1 implies

(4.16) $$\mu' \leqq \nu.$$

We conjecture

Conjecture. *Let μ be the minimal index at the origin (Definition 2.1). Let μ'*

be given as above. Then

$$\mu = \mu'. \tag{4.17}$$

In case that $\mu' = 0$ we have that $\kappa = \kappa(0)$ and the conjecture follows from Wasow's result. Suppose that $\mu' = 1$. Thus $A(\varepsilon)$ satisfies the conditions of Theorem 4.2 (see Introduction). It is easy to show that $A(\varepsilon)$ given in (1.5) fulfills the conditions of Theorem 4.2. Therefore the example (1.5) manifests that $\mu > 0$. Thus, indeed, Theorem 4.2 establishes the equality (4.17) in case that $\mu' = 1$. To prove Theorem 4.2 we need an auxiliary lemma.

LEMMA 4.2. *Let X and Y be $m \times m$ matrices. Assume that $r(X) = k$. Consider the subspace $\mathcal{U}$ of all vectors x of the form*

$$Xx = 0, \qquad \xi' Yx = 0, \qquad \xi' X = 0, \tag{4.18}$$

for all possible ξ. Assume that

$$\dim \mathcal{U} = m - k' \quad (\leqq m - k). \tag{4.19}$$

Then all $k' \times k'$ minors of $X + \varepsilon Y$ are of the form $\varepsilon^{k'-k} O(1)$. Moreover there exists an $k' \times k'$ minor $Q(\varepsilon)$ of $X + \varepsilon Y$ such that

$$|Q(\varepsilon)| = b\varepsilon^{k'-k}(1 + \varepsilon O(1)), \qquad b \neq 0. \tag{4.20}$$

PROOF. From Lemma 2.1 it follows that any $k' \times k'$ minor of $X + \varepsilon Y$ is of the form $\varepsilon^{k'-k} O(1)$. Suppose that (4.20) does not hold. Thus all $k' \times k'$ minors of $X + \varepsilon Y$ are of the form $\varepsilon^{k'-k+1} O(1)$. Let S_1, S_2 be two nonsingular matrices. Applying the Cauchy–Bihet formula we deduce that all $k' \times k'$ minors of $S_1 X S_2 + \varepsilon S_1 Y S_2$ are of the form $\varepsilon^{k'-k+1} O(1)$. We establish the lemma by showing that the above conclusion fails for some choice of nonsingular S_1 and S_2. Let

$$X_1 = S_1 X S_2, \qquad Y_1 = S_1 Y S_2. \tag{4.21}$$

We can choose S_1 and S_2 such that

$$X_1 = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}, \qquad Y_1 = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix}, \qquad Y_{22} = \begin{pmatrix} I_l & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.22}$$

Here X_1 and Y_1 are partitioned in the same manner and I_j is the $j \times j$ identity matrix. Clearly (4.18)–(4.19) holds if we replace X and Y by X_1 and Y_1. However in that case we immediately deduce that $m - k' = m - k - l$. Consider $k' \times k'$ minor $Q(\varepsilon)$ of $X_1 + \varepsilon Y_1$ based on the first k' rows and columns. Applying the Laplace expansion to the last l rows of $Q(\varepsilon)$ we deduce straightforward (4.20) with $b = 1$. This establishes the lemma.

PROOF OF THEOREM 4.2. Consider the expansion $\hat{A}(\varepsilon)$ given by (2.30). Let

$$X = I \otimes A_0 - A_0' \otimes I, \qquad Y = I \otimes A_1 - A_1' \otimes I. \tag{4.23}$$

So $r(X) = n^2 - \kappa(0)$ and $\dim \mathcal{U} = \kappa$. Thus according to Lemma 4.2 the conditions of Theorem 2.3 are satisfied so $\mu \leqq 1$. This in return is equivalent to (3.2)–(3.3). That is $A_1 \sim B_1(A_0)$.

We conclude this section with a different formulation of the system (4.1). Let $A_0, \cdots, A_{j-1}$ be $n \times n$ matrices. Define $C(A_0, \cdots, A_{j-1})$ to be $nj \times nj$ matrix which is block upper triangular:

$$C(A_0, \cdots, A_{j-1}) = (C_{pq})_1^j, \; C_{pq} = 0 \quad \text{for } q < p, \qquad C_{pq} = A_{q-p} \quad \text{for } q \geqq p \tag{4.24}$$

DEFINITION 4.2. Let $A_0, B_0, \cdots, A_{j-1}, B_{j-1}$ be given $n \times n$ matrices. The matrices $C(A_0, \cdots, A_{j-1})$ and $C(B_0, \cdots, B_{j-1})$ are called strongly similar if there exist $n \times n$ matrices $R_0, \cdots, R_{j-1}$ satisfying

$$\begin{gathered} C(A_0, \cdots, A_{j-1})C(R_0, \cdots, R_{j-1}) = C(R_0, \cdots, R_{j-1})C(B_0, \cdots, B_{j-1}), \\ \text{where } |R_0| \neq 0. \end{gathered} \tag{4.25}$$

As

$$|C(R_0, \cdots, R_{j-1})| = |R_0|^j \tag{4.26}$$

the assumption that $|R_0| \neq 0$ implies in particular that $C(A_0, \cdots, A_{j-1})$ is similar to $C(B_0, \cdots, B_{j-1})$. Now the system (4.1) for $k = 0, \cdots, j-1$ is equivalent to one matrix equation (4.25).

THEOREM 4.3. *Let $A(\varepsilon)$ and $B(\varepsilon)$ be $n \times n$ matrices analytic in ε at the origin. Then (1.10)–(1.11) are satisfied if and only if $C(A_0, \cdots, A_\omega)$ and $C(B_0, \cdots, B_\omega)$ are strongly similar. In particular if $A(\varepsilon) \sim_a B(\varepsilon)$ then $C(A_0, \cdots, A_\omega)$ and $C(B_0, \cdots, B_\omega)$ are similar for any $\omega \geqq 0$.*

It is left to show that the notion of strong similarity is indeed stronger than the similarity notion. Choose

$$A_0 = B_0 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B_1 = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}. \tag{4.27}$$

According to Theorem 3.2 $C(A_0, A_1)$ is strongly similar to $C(A_0, B_1)$ if and only if

$$a_{11} + a_{22} = c_{11} + c_{22}, \qquad a_{21} = c_{21}. \tag{4.28}$$

On the other hand if $a_{21} \neq 0$ then $C(A_0, A_1)$ has only one linearly independent eigenvector. Thus if $a_{21} \neq 0$, $C(A_0, A_1)$ is similar to

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Therefore if $a_{21} \neq 0$ and $c_{21} \neq 0$, $C(A_0, A_1)$ and $C(A_0, B_1)$ are similar.

5. Observations and remarks

We observe that the general problem stated in terms of the equations (4.1) for $k = 0, 1, \cdots, \omega$ is in fact of the same degree of complexity as Problem 3.1 (i.e. $\omega = 1$). More precisely we have

THEOREM 5.1. *Let Z be $kn \times kn$, a block diagonal matrix of the form*

$$Z = \operatorname{diag}\{H, \cdots, H\}, \qquad H = (\delta_{(i+1)j})_1^n. \tag{5.1}$$

Let X and Y be $kn \times kn$ block matrices

$$X = (X_{pq})_1^k, \quad X_{pq} = (x_{ij}^{(pq)})_1^n, \quad Y = (Y_{pq})_1^k, \quad Y_{pq} = (y_{ij}^{(pq)})_1^n. \tag{5.2}$$

Define

$$\begin{gathered} A_r = (a_{pq}^{(r)})_1^k, \qquad B_r = (b_{pq}^{(r)})_1^k, \\ a_{pq}^{(r)} = \sum_{i=1}^{r+1} x_{(n-r+i-1)i}^{(pq)}, \quad b_{pq}^{(r)} = \sum_{i=1}^{r+1} y_{(n-r+i-1)i}^{(pq)}, \qquad r = 0, \cdots, n-1. \end{gathered} \tag{5.3}$$

Then X is conjugated to Y with respect to Z if and only if $C(A_0, \cdots, A_{n-1})$ is strongly similar to $C(B_0, \cdots, B_{n-1})$.

To prove the theorem we need the following lemma.

LEMMA 5.1. *Let X be an $kn \times kn$ block matrix given by (5.2). Assume furthermore that each X_{pq} matrix is an upper triangular matrix. Then*

$$|X| = \prod_{r=1}^{n} |(x_{rr}^{(pq)})_{p,q=1}^k|. \tag{5.4}$$

PROOF. Expand X be the rows $n, 2n, \cdots, kn$. Obviously the only $k \times k$ non-vanishing minor which consists of $n, 2n, \cdots, kn$ rows is the minor composed of the columns $n, 2n, \cdots, kn$ of X. This minor is equal to $|(x_{nn}^{(pq)})_1^k|$. Now the lemma follows by induction.

PROOF OF THEOREM 5.1. According to Lemma 3.2 if P commutes with Z then P has the following form:

$$(5.5)\qquad P=(P_{pq})_1^k,\quad P_{pq}=\sum_{i=0}^{n-1} r_{pq}^{(i)}H^i,\quad R_i=(r_{pq}^{(i)})_1^k,\qquad i=0,\cdots,n-1.$$

Here $R_0,\cdots,R_{n-1}$ are arbitrary $k\times k$ matrices. According to Lemma 5.1

$$(5.6)\qquad P=|R_0|^n.$$

The subspace of all commuting matrices with Z is spanned by k^2n linearly independent matrices

$$(5.7)\qquad V_{pqi}=(V_{\alpha\beta}^{(pqi)})_1^k,\quad V_{\alpha\beta}^{(pqi)}=\delta_{\alpha p}\delta_{\beta q}H^i,\quad \alpha,\beta,p,q=1,\cdots,k,\ i=0,\cdots,n-1.$$

According to Lemma 3.1, P satisfies (3.5) for some Q if and only if

$$(5.8)\qquad \mathrm{tr}[V_{pqi}(XP-PY)]=0,\qquad p,q=1,\cdots,k,\quad i=0,\cdots,n-1.$$

Now

$$(5.9)\qquad \begin{aligned}\mathrm{tr}[V_{pqi}(XP-PY)]&=\mathrm{tr}\left[\sum_{j=1}^{k}(X_{qj}P_{jp}-P_{qj}Y_{jp})H^i\right]\\ &=\sum_{m=0}^{n-i-1}[r_{jp}^{(m)}\mathrm{tr}(X_{qj}H^{m+i})-r_{qj}^{(m)}\mathrm{tr}(Y_{jp}H^{m+i})].\end{aligned}$$

Note that (5.3) is equivalent to

$$(5.10)\qquad a_{pq}^{(r)}=\mathrm{tr}(X_{pq}H^{n-r-1}),\qquad b_{pq}^{(r)}=\mathrm{tr}(Y_{pq}H^{n-r-1}).$$

Thus (5.8) for $p,q=1,\cdots,k$ reduces to

$$(5.11)\qquad \sum_{m=0}^{n-i-1}(A_{n-m-i-1}R_m-R_mB_{n-m-i-1})=0,\qquad i=0,\cdots,n-1.$$

That is, we have the equalities (4.1) for $\omega=n-1$. The assumption that P is non-singular together with (5.6) yields that R_0 is non-singular. So $C(A_0,\cdots,A_{n-1})$ is strongly similar to $C(B_0,\cdots,B_{n-1})$. The proof of the theorem is concluded.

So if Z is of the form $\mathrm{diag}\{H,H\}$ then Problem 3.1 is reducible to the equalities (4.1) with $\omega=n-1$ where all matrices are 2×2. This in principle should not be difficult.

We conclude our paper with the following remarks about pointwise similarity of $A(\varepsilon)$ and $B(\varepsilon)$ in D_r for a small r. Obviously if $A(\varepsilon)\sim_p B(\varepsilon)$ then they must have the same characteristic polynomial

$$\lambda^n + \sum_{j=1}^{n} a_j(\varepsilon)\lambda^{n-j} = 0. \tag{5.12}$$

Moreover there exists $r' > 0$ such that for $0 < |\varepsilon| < r'$ the invariant polynomials of $A(\varepsilon)$ and $B(\varepsilon)$ are also analytic functions in ε. Therefore the elementary divisors $\varphi_1(\lambda, \varepsilon), \cdots, \varphi_p(\lambda, \varepsilon)$ and $\psi_1(\lambda, \varepsilon), \cdots, \psi_q(\lambda, \varepsilon)$ of $A(\varepsilon)$ and $B(\varepsilon)$ respectively are analytic in ε for $0 < |\varepsilon| < r''$. This in particular means that in this region the degrees of the elementary divisors are constant. So if $A(\varepsilon)$ and $B(\varepsilon)$ have the same characteristic polynomial (5.12) and are similar at $0 < |\varepsilon_0| < r''$ they must be pointwise similar for $0 < |\varepsilon| \leqq r''$! Thus if in addition $A(0) \sim B(0)$ we have that $A(\varepsilon) \sim_p B(\varepsilon)$, $\varepsilon \in D_{r''}$.

Added in proof. Recently, in the paper *Analytic similarity of matrices*, the author verified the conjecture stated here.

References

1. F. R. Gantmacher, *The Theory of Matrices*, Vol. I, Chelsea, New York, 1964.
2. M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Prindle, Weber and Schmidt, 1964.
3. W. Wasow, *On holomorphically similar matrices*, J. Math. Anal. Appl. 4 (1962), 202–206.
4. W. Wasow, *Topics in the theory of linear ordinary differential equations having singularities with respect to a parameter*, IRMA, Universite Louis Pasteur, Strasbourg, 1978.

Mathematics Research Center
University of Wisconsin — Madison
Wisconsin, U.S.A.

On leave from the Hebrew University of Jerusalem.