

AD-A068 866

WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER
ON POINTWISE AND ANALYTIC SIMILARITY OF MATRICES.(U)
FEB 79 S FRIEDLAND
MRC-TSR-1928

F/G 12/1

DAA629-75-C-0024
NL

UNCLASSIFIED

| OF |

AD
A068866



END
DATE
FILMED

7-79

DDC

LEVEL

2NW

AD A068866

MRC Technical Summary Report #1928

ON POINTWISE AND ANALYTIC SIMILARITY
OF MATRICES

Shmuel Friedland

Mathematics Research Center
University of Wisconsin-Madison
610 Walnut Street
Madison, Wisconsin 53706

February 1979

(Received January 25, 1979)

*See back page
for 1473*

DDC
RECEIVED
MAY 23 1979
C

DDC FILE COPY

Approved for public release
Distribution unlimited

Sponsored by

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

ON POINTWISE AND ANALYTIC SIMILARITY OF MATRICES

Shmuel Friedland

Technical Summary Report #1928
February 1979

ABSTRACT

Let $A(\epsilon)$ and $B(\epsilon)$ be complex valued matrices analytic in ϵ at the origin. $A(\epsilon) \underset{p}{\sim} B(\epsilon)$ if $A(\epsilon)$ is similar to $B(\epsilon)$ for any $|\epsilon| < r$, $A(\epsilon) \underset{a}{\sim} B(\epsilon)$ if $B(\epsilon) = T(\epsilon)A(\epsilon)T^{-1}(\epsilon)$ and $T(\epsilon)$ is analytic and $|T(\epsilon)| \neq 0$ for $|\epsilon| < r$. In this paper we find a necessary and sufficient condition on $A(\epsilon)$ and $B(\epsilon)$ such that $A(\epsilon) \underset{a}{\sim} B(\epsilon)$ provided that $A(\epsilon) \underset{p}{\sim} B(\epsilon)$. 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.

AMS(MOS) Subject Classifications - 15A21, 15A54, 34A25, 34E99

Key Words - Analytic similarity of matrices, pointwise similarity of matrices, conjugacy of matrices

Work Unit Number 1 - Applied Analysis

SIGNIFICANCE AND EXPLANATION

The matrix problem considered in this paper arises when studying systems of ordinary differential equations in "boundary-layer" situations.

A simple example is the behavior of solutions of $\epsilon y'' + py' + qy = 0$ (*) as $\epsilon \rightarrow 0$. In the general situation, a system of first order equations is considered, $A(\epsilon)y' + Cy = 0$, where A, B are $n \times n$ matrices. The first step is to simplify the system using a similarity transformation, i.e. we set $A = (T^{-1})z$ and multiply through by T , replacing the system by

$TA(\epsilon)(T^{-1})x' + TC(T^{-1})x = 0$.

The matrix T is chosen so as to simplify the coefficient of the derivative.

A matrix $A(\epsilon)$ is said to be analytic in ϵ at the origin if $A(\epsilon)$ can be expanded in a power series in ϵ . The following problem arises when classifying the various kinds of singular behavior of solutions of (*) for small ϵ . Suppose that $B(\epsilon) = T(\epsilon)A(\epsilon)T^{-1}(\epsilon)$ where $A(\epsilon), B(\epsilon)$ are analytic in ϵ at the origin. What conditions must be imposed on $A(\epsilon)$ and $B(\epsilon)$ so that $T(\epsilon)$ is analytic in ϵ at the origin and $T(0)$ is nonsingular? This is the question answered in this paper.

ACCESSION for	
NTIS	White Section ☒
DDC	Buff Section ☐
UNANNOUNCED	☐
JUSTIFICATION	
BY	
DISTRIBUTION/AVAILABILITY CODES	
Dist	SPECIAL
A	

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

ON POINTWISE AND ANALYTIC SIMILARITY OF MATRICES

Shmuel Friedland

1. INTRODUCTION.

Let $A(\epsilon)$ and $B(\epsilon)$ be $n \times n$ complex valued matrices analytic in the 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

$$(1.1) \quad A(\epsilon) = \sum_{k=0}^{\infty} A_k \epsilon^k, B(\epsilon) = \sum_{k=0}^{\infty} B_k \epsilon^k, A_k, B_k \in M_n(\mathbb{C})$$

which converge in D_r . One says that $A(\epsilon)$ and $B(\epsilon)$ are pointwise similar in D_r (denote it by $A(\epsilon) \sim_p B(\epsilon)$) if $A(\epsilon)$ and $B(\epsilon)$ are similar for any $\epsilon \in D_r$. $A(\epsilon)$ and $B(\epsilon)$ are said to be analytically similar in D_r , (denote it by $A(\epsilon) \sim_a B(\epsilon)$) if there exists $T(\epsilon)$

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

such that

$$(1.3) \quad |T(\epsilon)| \neq 0 \quad \text{for } |\epsilon| < r'$$

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

$$(1.4) \quad B(\epsilon) = T(\epsilon)A(\epsilon)T^{-1}(\epsilon).$$

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

Problem 1.1. (Wasow [4]) Assume that $A(\epsilon) \sim_p B(\epsilon)$ in D_r . What other conditions should $A(\epsilon)$ and $B(\epsilon)$ satisfy in order that $A(\epsilon) \sim_a B(\epsilon)$ in D_r , for some $0 < r' \leq r$?

Consider the following example

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

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

Otherwise

$$(1.6) \quad \begin{aligned} A_1(\epsilon) &= \epsilon^{-1}(A(\epsilon) - I) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tilde{a} \\ B_1(\epsilon) &= \epsilon^{-1}(B(\epsilon) - I) = \begin{pmatrix} 0 & \epsilon \\ 0 & 0 \end{pmatrix}, \quad |z| < r'. \end{aligned}$$

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

$$(1.7) \quad A(\epsilon)X - XA(\epsilon) = 0.$$

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

$$(1.8) \quad \kappa(\epsilon) = \kappa \quad 0 < |\epsilon| < \rho.$$

Wasow's Condition [3]. Assume that

$$(1.9) \quad \kappa(0) = \kappa.$$

Then $A(\epsilon) \sim_a B(\epsilon)$ in D_r , if and only if $A(\epsilon) \sim_p B(\epsilon)$ 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 $A(\epsilon)$ and $B(\epsilon)$ be $n \times n$ matrices analytic in ϵ for $|\epsilon| < r$.

There exists a non-negative integer ω depending only on $A(\epsilon)$ such that

$A(\epsilon) \sim_a B(\epsilon)$ for $\epsilon \in D_{r'}, (r' > 0)$ if and only if $A(\epsilon) \sim_p B(\epsilon)$ for $\epsilon \in D_r$, and
there exists $R(\epsilon)$ of the form

$$(1.10) \quad R(\epsilon) = \sum_{k=0}^{\omega} R_k \epsilon^k, \quad R_k \in M_n(\mathbb{C}), \quad |R_0| \neq 0$$

such that

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

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 $A(\epsilon)$ which do not satisfy the Wasow condition.

Theorem 4.2. Let $A(\epsilon)$ 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

$$(1.12) \quad R_0 A_0 = A_0 R_0, \text{tr}[V(R_0 A_1 - A_1 R_0)] = 0$$

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

$$(1.13) \quad P B_1 - A_1 P = A_0 R - R A_0$$

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

$$(1.14) \quad B_0 = A_0.$$

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

We state a conjecture which determine 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

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

Now

$$(2.2) \quad A(\epsilon) R_0 - R_0 B(\epsilon) = o(1),$$

as we claimed.

Suppose now that the Wasow condition fails. That is

$$(2.3) \quad \kappa < \kappa(0).$$

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

$$(2.4) \quad \hat{A}(\epsilon) \hat{X} = 0.$$

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

$$(2.5) \quad \begin{aligned} \hat{A}(\epsilon) &= (\hat{a}_{(i,j),(p,q)}(\epsilon)), \quad \hat{X} = (x_{(i,j)}) = (\text{vector}), \\ \hat{a}_{(i,j),(p,q)} &= a_{ip} \delta_{jq} - \delta_{ip} a_{jq}, \quad i, j, p, q = 1, \dots, n. \end{aligned}$$

Using the tensor product one can write

$$(2.6) \quad \hat{A}(\epsilon) = A(\epsilon) \otimes I - I \otimes A(\epsilon)$$

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

$$(2.7) \quad |P(\epsilon)| = a \epsilon^s (1 + o(1)), \quad a \neq 0.$$

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

$$(2.8) \quad \eta = n^2 - \kappa.$$

We claim that if one can satisfy the conditions (1.10) and (1.11) with $\omega = s$ then

$A(\epsilon) \sim_a B(\epsilon)$. Indeed, assume that $\omega = s$ and (1.10) and (1.11) holds. Since $|R_0| \neq 0$ there exists $0 < r' \leq r$ such that $R(\epsilon)^{-1}$ exists for $|\epsilon| < r'$. Let

$$(2.9) \quad C(\epsilon) = R(\epsilon) B(\epsilon) R(\epsilon)^{-1}.$$

Clearly it is enough to show that $A(\epsilon) \approx C(\epsilon)$. Also $A(\epsilon) \approx C(\epsilon)$. Consider the system

$$(2.10) \quad A(\epsilon)Y - YC(\epsilon) = 0.$$

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

$$(2.11) \quad F(\epsilon)\hat{Y} = 0.$$

In tensor notation

$$(2.12) \quad F(\epsilon) = A(\epsilon) \otimes I - I \otimes C(\epsilon).$$

According to our assumptions

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

So

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

Consider the submatrix $P(\epsilon)$ of $\hat{A}(\epsilon)$. Assume that the n rows of $P(\epsilon)$ form the set

$J \subset N \times N$ ($N = \{1, 2, \dots, n\}$) and the n columns of $P(\epsilon)$ form the set $K \subset N \times N$.

Look at the corresponding submatrix $Q(\epsilon)$ of $F(\epsilon)$ which is formed by the rows J

and the columns K . From (2.14) and (2.7) it follows that

$$(2.15) \quad |Q(\epsilon)| = a\epsilon^s(1 + o(1)).$$

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

$$(2.16) \quad y_{ij}(\epsilon) = \delta_{ij} \quad \text{if } (i, j) \notin K$$

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

$$(2.17) \quad Y(0) = I.$$

Indeed consider the unique solution $X(\epsilon)$ of (1.7) satisfying the condition (2.16).

Clearly $X(\epsilon) = I$ 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

$$(2.18) \quad Y(\epsilon) = (1 + o(1))X(\epsilon) + o(1).$$

This establishes (2.17) and the analyticity of Y around the neighborhood of the origin. So then exists $0 < r'' \leq r'$ such that $Y(\epsilon)$ and $Y(\epsilon)^{-1}$ holomorphic in

$|\epsilon| < r^n$. This proves the existence of ω depending only on $A(\epsilon)$ such that (1.10), (1.11) together with the assumptions $A(\epsilon) \underset{p}{\sim} B(\epsilon)$ at the origin imply that $A(\epsilon) \underset{a}{\sim} B(\epsilon)$ at $\epsilon = 0$. Vice versa, if $A(\epsilon) \underset{a}{\sim} B(\epsilon)$ for $\epsilon \in D_r$, then $A(\epsilon) \underset{p}{\sim} B(\epsilon)$ for $\epsilon \in D_r$, and (1.10) and (1.11) hold for any integer n . The proof of theorem is completed.

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

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 n be given by (2.8) and consider all non-zero $n \times n$ minors of $A(\epsilon) \otimes I - I \otimes A(\epsilon)$ which must be of the form (2.7). Let v be the minimum of all possible exponents s appearing in (2.7). Then

$$(2.19) \quad \mu \leq v.$$

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

Theorem 2.3. Let $A(\epsilon)$ satisfy the assumptions of Theorem 2.1. Assume that the Wasow condition fails (i.e. (2.3) holds). Suppose that v given in Theorem 2.2 equals to

$$(2.20) \quad v = \kappa(0) = \kappa.$$

Then the minimal index of $A(\epsilon)$ 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 $k (< n)$. Then for any $n \times n$ matrix Y and analytic valued $n \times n$ matrix $Z(\epsilon)$ ($|\epsilon| < r$) the following relations hold

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

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

Proof. Let $A(\epsilon) = (a_{ij}(\epsilon))_1^n$ be an analytic valued matrix at $\epsilon = 0$. Let

$r = (r_1, \dots, r_n)$ be a vector with non-negative integer coordinates. As usual denote

$|r| = \sum_{i=1}^n r_i$. By $(a_{ij}^{(r_i)}(\epsilon))_1^n$ denote the matrix whose i -th row is the r_i -th derivative of $A(\epsilon)$. From the standard formula of the derivative of the determinant we deduce

$$(2.23) \quad \frac{d^p}{d\epsilon^p} |A(\epsilon)| = \sum_{|r|=p} \frac{p!}{r_1! \dots r_n!} |(a_{ij}^{(r_i)}(\epsilon))_1^n|$$

Put

$$(2.24) \quad A(\epsilon) = X + \epsilon Y + \epsilon^2 Z(\epsilon)$$

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

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

Let $G \begin{pmatrix} i_1, \dots, i_q \\ j_1, \dots, j_q \end{pmatrix}$ be a $q \times q$ minor of G composed of $i_1, \dots, i_q$ rows and $j_1, \dots, j_q$ columns of G . In view of (2.25) we have

$$(2.26) \quad G \begin{pmatrix} i_1, \dots, i_q \\ j_1, \dots, j_q \end{pmatrix} = X \begin{pmatrix} i_1, \dots, i_q \\ j_1, \dots, j_q \end{pmatrix}, \quad 1 \leq i_1, \dots, i_k \leq n.$$

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

$$(2.27) \quad \left. \frac{d^p}{d\epsilon^p} |A(\epsilon)| \right|_{\epsilon=0} = 0, \quad p = 0, \dots, n - k - 1.$$

Assume now that $p = n - k$. Again if $q \geq k + 1$, $|G| = 0$. So we are left with the case where $q = k$. That is, there exist $1 \leq i'_1 < i'_2 < \dots < i'_{n-k} \leq n$ such that

$$(2.28) \quad r_{i'_1} = \dots = r_{i'_{n-k}} = 1.$$

In this case G is composed of $i'_1, \dots, i'_{n-k}$ rows of Y and $i_1, \dots, i_k$ rows of X . Therefore we showed

$$(2.29) \quad \left. \frac{d^{n-k}}{d\epsilon^{n-k}} |A(\epsilon)| \right|_{\epsilon=0} = \left. \frac{d^{n-k}}{d\epsilon^{n-k}} |X + \epsilon Y| \right|_{\epsilon=0}.$$

This verifies (2.20) and (2.21).

Proof of Theorem 2.3. Assume that $B(\epsilon) \approx_p A(\epsilon)$ for $\epsilon \in D_r$. Suppose that (1.10) and (1.11) holds for $\omega = 1$. We claim that $B(\epsilon) \approx_a A(\epsilon)$ for $\epsilon \in D_r$, 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 v we may assume that s given in (2.7) equals to v . From (2.6), (2.12) and the equality $\omega = 1$ we get

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

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

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

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 $\hat{A}(\epsilon)$ ($F(\epsilon)$) from the set K and a column which is a linear combination of the columns of $\hat{A}(\epsilon)$ ($F(\epsilon)$) which do not belong to K . Clearly the rank of such a minor at $\epsilon = 0$ is at most $\eta(0)$. Using (2.31) and (2.22) we obtain that the difference between the corresponding minors of $\hat{A}(\epsilon)$ and $F(\epsilon)$ is at least of the form $\epsilon^{\eta - \eta(0) + 1} O(1)$, i.e. $\epsilon^{v+1} O(1)$. Dividing the minors of $\hat{A}(\epsilon)$ by $|P(\epsilon)|$ and the minors of $F(\epsilon)$ by $|Q(\epsilon)|$ from (2.7) and (2.15) we deduce (2.18). The proof of the theorem is completed.

3. THE CASE $\omega = 1$.

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

$$(3.1) \quad A_0 = B_0.$$

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

$$(3.2) \quad A_0 R_0 - R_0 A_0 = 0, \quad |R_0| \neq 0,$$

$$(3.3) \quad A_0 R_1 + A_1 R_0 - R_1 A_0 - R_0 B_1 = 0.$$

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

$$(3.4) \quad ZP - PZ = 0,$$

such that

$$(3.5) \quad XP - PY = ZQ - QZ$$

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

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

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

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

$$(3.6) \quad U = ZQ - QZ$$

for some Q , if and only if

$$(3.7) \quad \text{tr}(VU) = 0$$

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

Proof. Clearly if V commutes with Z then

$$(3.8) \quad \text{tr}(VU) = \text{tr}(VZQ - VQZ) = \text{tr}[Z(VQ) - VQZ] = \text{tr}[(VQ)Z - VQZ] = 0.$$

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} , $i, j = 1, \dots, n$ ($Q = (q_{ij})_1^n$). In tensor form (3.6) is given as

$$(3.9) \quad (Z \otimes I - I \otimes Z) \hat{Q} = \hat{U}$$

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

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

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

Now (3.11) means that

$$(3.12) \quad Z^T W - W Z^T = 0.$$

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

Let $v_1, \dots, v_k$ form a basis for the subspace of all matrices in $M_n(\mathbb{C})$ which commute with Z . Thus any P which satisfies (3.4) is of the form

$$(3.13) \quad P = \sum_{i=1}^k v_i v_i^T.$$

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

$$(3.14) \quad \text{tr}[v_j (XP - PY)] = 0, \quad j = 1, \dots, k.$$

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

In principle this can be done by verifying a finite number of conditions.

Indeed let

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

Thus 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

$$(3.16) \quad v_i = 0, 1, \dots, n, \quad i = 1, \dots, k.$$

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

$$(3.17) \quad F(tv_1, \dots, tv_k) = t^n F(v_1, \dots, v_k).$$

Next we observe that

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

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

$$(3.19) \quad Z = \text{diag}\{J_1, \dots, J_u\}, \quad J_k = \lambda_k I_k + H_k, \quad \dim J_k = n_k, \quad k = 1, \dots, u.$$

Here I_k is the identity matrix and H_k 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 $Z \in M_n(\mathbb{C})$ be a matrix given by (3.19). Then a block matrix

$P = (p_{\alpha\beta})_{\alpha, \beta=1}^u \in M_n(\mathbb{C})$ commutes with Z if and only if the blocks

$p_{\alpha\beta} = (p_{ij}^{(\alpha\beta)})_{i=1, \dots, n_\alpha, j=1, \dots, n_\beta}$ satisfy the following conditions

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

$$(3.21) \quad \text{if } \lambda_\alpha = \lambda_\beta \quad \begin{cases} p_{ij}^{(\alpha\beta)} = 0 & \text{for } j < i + n_\beta - \min(n_\alpha, n_\beta), \\ p_{ij}^{(\alpha\beta)} = p_{(i+1)(j+1)}^{(\alpha\beta)} & \text{for } j \geq i + n_\beta - \min(n_\alpha, n_\beta). \end{cases}$$

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

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

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

$$(3.23) \quad 0 = \text{tr} \left\{ V_{\alpha\beta} \left[\sum_{j=1}^u (X_{\beta j} P_{j\alpha} - P_{\beta j} Y_{j\alpha}) \right] \right\},$$

$$X = (X_{ij})_1^u, \quad Y = (Y_{ij})_1^u$$

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

$$(3.24) \quad Z = \text{diag}(Z_1, \dots, Z_v)$$

such that

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

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

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

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 $i = 1, \dots, v$.

Theorem 3.2. Assume that Z consists of one Jordan block

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

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

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

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

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

The assumption that P is a 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

$$(3.30) \quad 0 = \text{tr}(H^j X P - H^j P Y) = \text{tr}(X P H^j - Y H^j P) = \text{tr} \left[(X - Y) \left(\sum_{i=0}^{n-j-1} a_i H^{i+j} \right) \right], \quad j = n-1, \dots, 0.$$

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

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

Assume that we already proved (3.31) for $j = n-1, \dots, k$. 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 $j = n-1, \dots, 0$. 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

$$(4.1) \quad A_0 R_k - R_k A_0 = \sum_{i=1}^k (R_{k-i} B_i - A_i R_{k-i}),$$

for $k = 0, 1, \dots, \omega$, where we assumed the normalization $A_0 = B_0$. A sequence

$(R_0, \dots, R_j)$ is called a solution if $R_0, \dots, R_j$ satisfy (4.1) for $k = 0, 1, \dots, j$.

Denote by L_j the subspace of all solutions $(R_0, \dots, R_j)$ and by $L_{j,i}$ the subspace of the first i matrices $(R_0, \dots, R_i)$ in the solutions $(R_0, \dots, R_j)$ where $0 \leq i \leq j$.

Clearly

$$(4.2) \quad L_{j,i} \supseteq L_{j+1,i}.$$

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

$$(4.3) \quad \text{tr}[V \sum_{i=1}^{j+1} (R_{j+1-i} B_i - A_i R_{j+1-i})] = 0, \quad V A_0 = A_0 V,$$

for all V which commute with A_0 . Thus if we constructed L_j (4.3) determines $L_{j+1,j}$. Now by solving (4.1) for $k = j+1$ where $(R_0, \dots, R_j) \in L_{j+1,j}$ we obtain the subspace L_{j+1} . Thus if $A(\epsilon) \underset{p}{\sim} B(\epsilon)$ then $A(\epsilon) \underset{a}{\sim} B(\epsilon)$ if and only if $L_{v,0}$ contains a non-singular matrix (v -is given in Theorem 2.2).

Theorem 4.1. Assume that $A(\epsilon)$ and $B(\epsilon)$ are analytically similar at the origin.

Consider the system (4.1) for $k = 0, \dots, j$. Then

$$(4.4) \quad \dim L_{j,0} \geq \kappa$$

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

To prove this theorem we need the following lemma.

Lemma 4.1. Let $A(\epsilon)$ be complex valued matrix analytic in ϵ at the origin. Consider all complex valued matrices $X(\epsilon) = \sum_{k=0}^{\infty} X_k \epsilon^k$ 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

$$\dim U \leq \kappa.$$

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

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

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

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

$$(4.6) \quad |Q(\epsilon)| = a' \epsilon^{s'} (1 + o(1)), \quad a' \neq 0, \quad s' \geq 1.$$

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

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

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

$$(4.8) \quad \bar{X}^{(i)}(\epsilon) = X^{(i)}(\epsilon), \quad i = 1, \dots, \kappa - 1, \quad \bar{X}^{(\kappa)}(\epsilon) = \epsilon^{-1} \sum_{i=1}^{\kappa} \alpha_i X^{(i)}(\epsilon).$$

Let $\bar{H}(\epsilon)$ be the matrix composed of $\bar{X}^{(1)}, \dots, \bar{X}^{(\kappa)}$. Again if $r(\bar{H}(0)) = \kappa$ we are done. Otherwise considering the corresponding minor $\bar{Q}(\epsilon)$ which consists of the same

rows and columns as $Q(\epsilon)$, we easily deduce that

$$(4.9) \quad |\tilde{Q}(\epsilon)| = a' \epsilon^{s'-1} (1 + \epsilon O(1)) .$$

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

Proof of Theorem 4.1. Let $X(\epsilon)$ be an analytic solution of (1.7). Denote

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

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

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

We also have

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

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

$$(4.13) \quad \dim L_{v,0} = \kappa .$$

Assume that $R(\epsilon) = \sum_{k=0}^{\omega} R_k \epsilon^k$ satisfy (1.11). Here we do not demand that $|R_0| \neq 0$.

Moreover assume that $\omega = v$. Define $X(\epsilon)$ by the equation (4.10). From (1.11) and (1.4) we get

$$A(\epsilon) X(\epsilon) - X(\epsilon) A(\epsilon) = \epsilon^{v+1} O(1)$$

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

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

Definition 4.1. Consider the system of matrix equations

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

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

$$(4.15) \quad U_{\mu'-1} \supset U, U_{\mu'} = U$$

where U is given by Lemma 4.1.

Theorem 4.1 implies

$$(4.16) \quad \mu' \leq \nu.$$

We conjecture

Conjecture. Let μ be the minimal index at the origin (Definition 2.1). Let μ' be given as above. Then

$$(4.17) \quad \mu = \mu'.$$

In case that $\mu' = 0$ we have that $\kappa = \kappa(0)$ and the conjecture follows from Wasow's result. Suppose that $\mu' = 1$. Thus $A(\epsilon)$ satisfies the conditions of Theorem 4.2 (see Introduction). It is easy to show that $A(\epsilon)$ 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 U of all vectors x of the form

$$(4.18) \quad Xx = 0, \xi^T Yx = 0, \xi^T X = 0,$$

for all possible ξ . Assume that

$$(4.19) \quad \dim U = m - k' \quad (\leq m - k).$$

Then all $k' \times k'$ minors of $X + \epsilon Y$ are of the form $\epsilon^{k'-k} 0(1)$. Moreover there exists an $k' \times k'$ minor $Q(\epsilon)$ of $X + \epsilon Y$ such that

$$(4.20) \quad |Q(\epsilon)| = b\epsilon^{k'-k}(1 + \epsilon 0(1)), \quad b \neq 0.$$

Proof. From Lemma 2.1 it follows that any $k' \times k'$ minor of $X + \epsilon Y$ is of the form $\epsilon^{k'-k} 0(1)$. Suppose that (4.20) does not hold. Thus all $k' \times k'$ minors of $X + \epsilon Y$ are of the form $\epsilon^{k'-k+1} 0(1)$. Let S_1, S_2 be two nonsingular matrices. Applying the Cauchy-Bihet formula we deduce that all $k' \times k'$ minors of $S_1 X S_2 + \epsilon S_1 Y S_2$ are of the form $\epsilon^{k'-k+1} 0(1)$. We establish the lemma by showing that the above conclusion fails for some choice of nonsingular S_1 and S_2 . Let

$$(4.21) \quad X_1 = S_1 X S_2, \quad Y_1 = S_1 Y S_2.$$

We can choose S_1 and S_2 such that

$$(4.22) \quad X_1 = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}, \quad Y_1 = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix}, \quad Y_{22} = \begin{pmatrix} I_\ell & 0 \\ 0 & 0 \end{pmatrix}.$$

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 $m - k' = m - k - \ell$. Consider $k' \times k'$ minor $Q(\epsilon)$ of $X_1 + \epsilon Y_1$ based on the first k' rows and columns. Applying the Laplace expansion to the last ℓ rows of $Q(\epsilon)$ we deduce straightforward (4.20) with $b = 1$. This establishes the lemma.

Proof of Theorem 4.2. Consider the expansion $\hat{A}(\epsilon)$ given by (2.30). Let

$$(4.23) \quad X = A_0 \otimes I - I \otimes A_0, \quad Y = A_1 \otimes I - I \otimes A_1.$$

So $r(X) = n^2 - \kappa(0)$ and $\dim U = \kappa$. Thus according to Lemma 4.2 the conditions of Theorem 2.3 are satisfied so $\mu \leq 1$. 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 $A_0, \dots, A_{j-1}$ be $n \times n$ matrices. Define $C(A_0, \dots, A_{j-1})$ to be $nj \times nj$ matrix which is block upper triangular

$$(4.24) \quad C(A_0, \dots, A_{j-1}) = (C_{pq})_1^j, \quad C_{pq} = 0 \text{ for } q < p, \quad C_{pq} = A_{q-p} \text{ for } q \geq p$$

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

$$(4.25) \quad C(A_0, \dots, A_{j-1}) C(R_0, \dots, R_{j-1}) = C(R_0, \dots, R_{j-1}) C(B_0, \dots, B_{j-1}), \quad \text{where } |R_0| \neq 0$$

As

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

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

Theorem 4.3. Let $A(\epsilon)$ and $B(\epsilon)$ be $n \times n$ matrices analytic in ϵ at the origin. Then (1.10) - (1.11) are satisfied if and only if $C(A_0, \dots, A_\omega)$ and $C(B_0, \dots, B_\omega)$ are strongly similar. In particular if $A(\epsilon) \succeq B(\epsilon)$ then $C(A_0, \dots, A_\omega)$ and $C(B_0, \dots, B_\omega)$ are similar for any $\omega \geq 0$.

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

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

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

$$(4.28) \quad a_{11} + a_{22} = c_{11} + c_{22}, \quad a_{21} = c_{21}.$$

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 $k = 0, 1, \dots, \omega$ 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

$$(5.1) \quad Z = \text{diag}\{H, \dots, H\}, \quad H = (\delta_{(i+1)j})_{1}^n$$

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

$$(5.2) \quad X = (x_{pq})_{1}^k, \quad x_{pq} = (x_{ij}^{(pq)})_{1}^n, \quad Y = (y_{pq})_{1}^k, \quad y_{pq} = (y_{ij}^{(pq)})_{1}^n.$$

Define

$$(5.3) \quad A_r = (a_{pq}^{(r)})_{1}^k, \quad B_r = (b_{pq}^{(r)})_{1}^k, \\ a_{pq}^{(r)} = \sum_{i=1}^{r+1} x_{(n-r+i-1)i}^{(pq)}, \quad b_{pq}^{(r)} = \sum_{i=1}^{r+1} y_{(n-r+i-1)i}^{(pq)}, \quad r = 0, \dots, n-1$$

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

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

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

Proof. Expand X by the rows $n, 2n, \dots, kn$. Obviously the only $k \times k$ non-vanishing minor which consists of $n, 2n, \dots, kn$ rows is the minor composed of the columns $n, 2n, \dots, kn$ of X . This minor is equal $|(x_{nn}^{(pq)})_{1}^k|$. 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

$$(5.5) \quad P = (p_{pq})_{1}^k, \quad p_{pq} = \sum_{i=0}^{n-1} r_{pq}^{(i)} H^i, \quad R_i = (r_{pq}^{(i)})_{1}^k, \quad i = 0, \dots, n-1.$$

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

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

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

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

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

$$(5.8) \quad \text{tr}[v_{pqi}(XP - PY)] = 0, \quad p, q = 1, \dots, k, \quad i = 0, \dots, n-1.$$

Now

$$(5.9) \quad \begin{aligned} \text{tr}[v_{pqi}(XP - PY)] &= \text{tr} \left[\sum_{j=1}^k (x_{qj}^p x_{jp} - y_{jp}^q y_{jq}) H^i \right] \\ &= \sum_{m=0}^{n-i-1} [r_m^{(jp)} \text{tr}(x_{qj} H^{m+i}) - r_m^{(qj)} \text{tr}(y_{jp} H^{m+i})] \end{aligned}$$

Note that (5.3) is equivalent to

$$(5.10) \quad a_{pq}^{(r)} = \text{tr}(x_{pq} H^{n-r-1}), \quad b_{pq}^{(r)} = \text{tr}(y_{pq} H^{n-r-1}).$$

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

$$(5.11) \quad \sum_{m=0}^{n-i-1} (A_{n-m-i-1} R_m - R_m B_{n-m-i-1}) = 0, \quad i = 0, \dots, n-1.$$

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 $C(A_0, \dots, A_{n-1})$ is strongly similar to $C(B_0, \dots, B_{n-1})$. The proof of the theorem is concluded.

So if Z is of the form $\text{diag}\{H, H\}$ 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 $A(\epsilon)$ and $B(\epsilon)$ in D_r for a small r . Obviously if $A(\epsilon) \sim_p B(\epsilon)$ then they must have the same characteristic polynomial

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

Moreover there exists $r' > 0$ such that for $0 < |\epsilon| < r'$ the invariant polynomials of $A(\epsilon)$ and $B(\epsilon)$ are also analytic functions in ϵ . Therefore the elementary divisors $\varphi_1(\lambda, \epsilon), \dots, \varphi_p(\lambda, \epsilon)$ and $\psi_1(\lambda, \epsilon), \dots, \psi_q(\lambda, \epsilon)$ of $A(\epsilon)$ and $B(\epsilon)$ respectively are analytic in ϵ for $0 < |\epsilon| < r''$. This in particular means that in this region the degrees of the elementary divisors are constant. So if $A(\epsilon)$ and $B(\epsilon)$ have the same characteristic polynomial (5.12) and are similar at $0 < |\epsilon_0| < r''$ they must be pointwise similar for $0 < |\epsilon| \leq r''$. So if $A(0) \sim B(0)$ we have that $A(\epsilon) \sim_p B(\epsilon)$, $\epsilon \in D_{r''}$.

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. and Appl. 4, 202-206 (1962).
- [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.

SF/ed

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 1928	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER 9/ Technical
4. TITLE (and Subtitle) ON POINTWISE AND ANALYTIC SIMILARITY OF MATRICES		5. TYPE OF REPORT & PERIOD COVERED Summary Report, no specific reporting period
7. AUTHOR(s) Shmuel Friedland		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of 610 Walnut Street Wisconsin Madison, Wisconsin 53706		8. CONTRACT OR GRANT NUMBER(s) 15/ DAAG29-75-C-0024
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 1 - Applied Analysis
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 12/ 27p.		12. REPORT DATE 11/ February 1979
		13. NUMBER OF PAGES 23
		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited. 14/ MRC-TSR-1928		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Analytic similarity of matrices, pointwise similarity of matrices, conjugacy of matrices		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Let $A(\epsilon)$ and $B(\epsilon)$ be complex valued matrices analytic in ϵ at the origin. $A(\epsilon) \sim B(\epsilon)$ if $A(\epsilon)$ is similar to $B(\epsilon)$ for any $ \epsilon < r, A(\epsilon) \sim_a B(\epsilon)$ if $B(\epsilon) = T(\epsilon)A(\epsilon)T^{-1}(\epsilon)$ and $T(\epsilon)$ is analytic and $ T(\epsilon) \neq 0$ for $ \epsilon < r$. In this paper we find a necessary and sufficient condition on $A(\epsilon)$ and $B(\epsilon)$ such that $A(\epsilon) \sim_a B(\epsilon)$ provided that $A(\epsilon) \sim B(\epsilon)$. 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.		