

On Simultaneously and Approximately Simultaneously Diagonalizable m -tuples of Matrices

Maria Isabel García-Planas

Departament de Matemàtiques, Universitat Politècnica de Catalunya, Barcelona, Spain

Article Info

Keywords: Diagonalization, Eigenvalues, Eigenvectors, Equivalence relation, Simultaneously diagonalization

2010 AMS: 15A21, 15A22

Received: 26 March 2019

Accepted: 17 May 2019

Available online: X XXXXXX XXXX

Abstract

In this paper, the problem of simultaneous diagonalization of m -tuples of n -order square complex matrices, is analyzed and some necessary and some necessary and sufficient conditions for this property to be fulfilled are presented. This study has an interest in its applications in different areas as for example in engineering and physical sciences. For example, they appear founding when we must give the instanton solution of Yang-Mills field presented in an octonion form, and it can be represented by triples of traceless matrices. In the case where the m -tuple does not simultaneously diagonalize, the possibility of to find near of the given m -tuple, an m -tuple that diagonalize simultaneously is studied.

1. Introduction

Let \mathfrak{M} be the manifold of m -tuples of n -order square complex matrices $T = (X_1, \dots, X_m)$ representing polynomial matrices $P_T(x) = X_1 + xX_2 + \dots + x^{m-1}X_m$ that appear in a natural way modeling tools in several research areas of applied mathematics, sciences and engineering, and in a special manner in systems theory ([1]-[3]). Studying control problems by means the polynomial matrix approach, the solution of these problems are reformulated in terms of polynomial matrix equations, where solutions are based on structural properties of the involved matrices, where the simultaneous diagonalization of each and every one of the matrices is a great advantage for solving the problem. The simultaneously diagonalization is related to sets of commuting matrices and it can be found some results (see [4], [5], for example). Among families of m -tuples of matrices, have some interest the families of traceless triples because the Lie algebra is related to gauge fields because they appear in the Lagrangian describing the dynamics of the field, then they are associated to 1-forms that take values on a certain Lie algebra. It is also of interest to note that triples of traceless matrices have some relevance for supergravity theories ([6]). Another application is found when we must give the instanton solution of Yang-Mills field can be presented in an octonion form, and it can be represented by triples of traceless matrices ([7]).

In the space of n -square complex matrices, it is well known that the subset of diagonalizable matrices is generic in the sense that this subset is an open and dense set, then any no diagonalizable matrix can be diagonalized by a small perturbation of its entries. This property cannot be generalized to the case of simultaneous diagonalization of an m -tuple of n -order complex square matrices. We are interested in analyzing the collection of m -tuples of matrices that simultaneously diagonalize and the collection that simultaneously diagonalize under small perturbations, some properties in this sense appear in [8].

The simultaneous diagonalization of two real symmetric matrices has long been of interest and largely studied [9]. In this paper, we generalize to the problem of deciding whether the elements of \mathfrak{M} can be simultaneously diagonalized, and in the case where the m -tuple does not simultaneously diagonalize, we study the possibility of to find near of the given m -tuple, an m -tuple that diagonalize simultaneously.

2. Simultaneous similarity of m -tuples of n -order matrices

Definition 2.1. Let $T = (X_1, \dots, X_m), T' = (Y_1, \dots, Y_m) \in \mathfrak{M}$ be two m -tuples of matrices. Then, T is simultaneous similar to T' if and only if there exists $P \in GL(n; \mathbb{R})$ such that

$$(Y_1, \dots, Y_m) = (PX_1P^{-1}, \dots, PX_mP^{-1}). \quad (2.1)$$

For simplicity, we will write $PTP' = T'$.

We are interested on the simultaneous diagonalizable m -tuples.

Definition 2.2. The m -tuples of matrices $T = (X_1, \dots, X_m) \in \mathfrak{M}$ is simultaneously diagonalizable if and only if there exist an equivalent m -tuple formed by diagonal matrices.

From definition we have

Corollary 2.3. Let $T = (X_1, \dots, X_m)$ be an m -tuple of square matrices. The m -tuple is simultaneous diagonalizable if and only if there exist diagonal matrices $D_i, i = 1, \dots, m$ and a invertible matrix P (the same matrix P for all i) such that

$$(X_i^t \otimes I_n - I_n \otimes D_i) \text{vec} P = 0, \forall 1 \leq i \leq m.$$

Remark 2.4. Let $A = (a_{ij})$ and B , the Kronecker product is defined as $A \otimes B = (a_{ij}B)$.

Proof. From $D_i = PX_iP^{-1}$ for all $i = 1, \dots, m$ we have $PX_i - D_iP = 0$, for all $i = 1, \dots, m$

Then, computing the Kronecker product and applying the vectorizing operator we deduce the result. □

Clearly, necessary conditions for simultaneous diagonalizable m -tuples are the following

Proposition 2.5. Let $T = (X_1, \dots, X_m)$ be a simultaneous diagonalizable m -tuple. Then all matrices X_i must be diagonalizable.

Obviously, the reciprocal is false

Example 2.6. Clearly, matrices $X_1 = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ and $X_2 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ are diagonalizable, but none of the matrices $P_i = Q_i^{-1}$ with $Q_1 = \begin{pmatrix} a & b \\ 0 & b \end{pmatrix}$ and $Q_2 = \begin{pmatrix} b & a \\ 0 & 0 \end{pmatrix}$ with $ab \neq 0$, diagonalizing X_1 can diagonalize X_2 .

Proposition 2.7. Let $T = (X_1, \dots, X_m)$ be a simultaneous diagonalizable m -tuple. Then $X_iX_j = X_jX_i$.

Proof. Let $T = (X_1, \dots, X_m)$ be a simultaneously diagonalizable m -tuple, then there exist $P \in GL(n; \mathbb{C})$ such that $PX_iP^{-1} = D_i$ for $i = 1, \dots, m$. So, taking into account that $D_iD_j = D_jD_i$, for all $i, j = 1, \dots, m$, we have $P^{-1}D_iPP^{-1}D_jP = P^{-1}D_jPP^{-1}D_iP$, that is to say $X_iX_j = X_jX_i$, for all $i, j = 1, \dots, m$. □

Theorem 2.8. Let $T = (X_1, \dots, X_m)$ be a m -tuple of commuting n -order square matrices and suppose that the matrix X_j for some j is diagonalizable with simple eigenvalues ($\lambda_k \neq \lambda_\ell$ for all $k \neq \ell, k, \ell = 1, \dots, n$). Then T is a m -tuple of simultaneously diagonalizable matrices

Proof. For simplicity we consider X_1 the diagonalizable matrix.

Let $\{v_1, \dots, v_n\}$ be a basis of eigenvectors corresponding to eigenvalues $\{\lambda_1, \dots, \lambda_n\}$ of X_j .

Let us consider $X_iX_1v_j$ for all $i = 1, \dots, m$ and $j = 1, \dots, n$.

$$\begin{aligned} X_iX_1v_j &= X_i\lambda_jv_j = \lambda_jX_iv_j \\ X_jX_iv_j &= \lambda_jX_iv_j \end{aligned}$$

So, if $X_iv_j \neq 0$ it is an eigenvector of X_1 of eigenvalue λ_j , but condition $\lambda_k \neq \lambda_\ell$ implies that $\dim \text{Ker}(X_1 - \lambda_jI) = 1$, then, $X_iv_j = \mu_iv_j$, that is to say v_j is an eigenvector for X_i of eigenvalue μ_i . If $X_iv_j = 0$ v_j is an eigenvector of X_i of eigenvalue equal zero. That is to say $\{v_1, \dots, v_n\}$ is a basis of eigenvectors for each $X_i, i = 1, \dots, m$ and T is m -tuple of simultaneous diagonalizable matrices with $P = (v_1^t \dots v_n^t)^{-1}$. □

Remark 2.9. The other matrices not necessary have simple eigenvalues.

Theorem 2.10. Let $T = (X_1, \dots, X_m)$ be an m -tuple of commuting and diagonalizable n -order square matrices. Then, they diagonalize simultaneously.

Proof. Let P_1 be an invertible matrix such that $D_1 = P_1X_1P^{-1} = \begin{pmatrix} D_1^1 & & \\ & \ddots & \\ & & D_1^r \end{pmatrix}$ with $D_i^1 = \lambda_i^1I \in M_{n_i}(\mathbb{C}), 1 \leq i \leq r$ and $n_1 + \dots + n_r = n$.

Let us consider $v_{1_1}, \dots, v_{n_1}, \dots, v_{1_r}, \dots, v_{n_r}$, the vector columns of P^{-1} , then

$$\begin{aligned} X_jX_\ell v_{i_\ell} &= X_j\lambda_\ell v_{i_\ell} = \lambda_\ell X_jv_{i_\ell} \\ X_jX_\ell v_{i_\ell} &= X_\ell X_jv_{i_\ell} \end{aligned}$$

Consequently $X_jv_{i_\ell}$ is an eigenvector of X_ℓ of eigenvalue λ_ℓ or $X_jv_{i_\ell} = 0$, in any case we have that $X_jv_{i_\ell} \in [v_{1_\ell}, \dots, v_{n_\ell}] = F_\ell$, consequently, the subspace F_ℓ is X_j invariant for all $1 \leq \ell \leq r$ and $1 \leq j \leq m$.

So, $P_1X_jP_1^{-1} = \begin{pmatrix} Y_1^j & & \\ & \ddots & \\ & & Y_r^j \end{pmatrix}$, for $2 \leq j \leq m$.

If all matrices Y_k^j are diagonal the proof is concluded, otherwise and taking into account that all matrices X_i diagonalize all submatrices Y_k^j diagonalize.

Consider $P_2 = \begin{pmatrix} P_2^1 & & \\ & \ddots & \\ & & P_1^r \end{pmatrix}$ where P_2^j diagonalizes Y_j^2 for $1 \leq j \leq r$.

Obviously P_2 diagonalizes D_1 :

$$\begin{pmatrix} P_2^1 & & \\ & \ddots & \\ & & P_2^r \end{pmatrix} \begin{pmatrix} D_1^1 & & \\ & \ddots & \\ & & D_1^r \end{pmatrix} \begin{pmatrix} P_2^1 & & \\ & \ddots & \\ & & P_2^r \end{pmatrix}^{-1} = \begin{pmatrix} P_2^1 D_1^1 (P_2^1)^{-1} & & \\ & \ddots & \\ & & P_2^r D_1^r (P_2^r)^{-1} \end{pmatrix} = \begin{pmatrix} D_1^1 & & \\ & \ddots & \\ & & D_1^r \end{pmatrix}$$

Then $P_2 P_1$ diagonalizes X_1 and X_2 , now partitioning the matrices $P_2^j Y_j^2 (P_2^j)^{-1}$ into blocks corresponding to the same eigenvalue (it is possible that different blocks Y_j^2 have common eigenvalues but we partition according to each block).

Now we consider $P_2 P_1 X_j (P_2 P_1)^{-1}$, if all matrices are diagonal the proof is concluded, otherwise we repeat the process with $P_2 P_1 X_3 (P_2 P_1)^{-1}$ taking into account the new partition in scalar matrices. The process ends at most when reaches to the last matrix. \square

After these results it is easy to obtain the following geometrical result.

Theorem 2.11. Let $T = (X_1, \dots, X_m)$ be an m -tuple of n -order square matrices and suppose that all matrices X_i are diagonalizable, then a necessary and sufficient condition for simultaneous diagonalization is there exist a basis $\{v_1, \dots, v_n\}$ of $v \in \mathbb{C}^n$ such that

$$v_j \in \bigcap_{i=1}^m \text{Ker}(X_i - \lambda_i^j I), \text{ where } \lambda_i^j \in \text{Spec} X_i = \{\lambda_1^i, \dots, \lambda_n^i\}$$

Corollary 2.12.

$$P = (v_1^t \quad \dots \quad v_n^t)^{-1}$$

verifies that $PX_i P^{-1} = D_i$

Example 2.13. Let $T = (X_1, X_2, X_3)$ be a triple with

$$X_1 = \begin{pmatrix} 5.5 & 2 & -3.5 \\ 3 & 3 & -3 \\ 4.5 & 2 & -2.5 \end{pmatrix}, X_2 = \begin{pmatrix} 3 & -2 & 2 \\ 1.5 & 4 & -1.5 \\ 0 & -2 & 5 \end{pmatrix}, X_3 = \begin{pmatrix} 15.5 & 10 & -13.5 \\ 3 & 7 & -3 \\ 10.5 & 10 & -8.5 \end{pmatrix}$$

$$\text{Spec} X_1 = \{1, 2, 3\}, \text{Spec} X_2 = \{3, 4, 5\}, \text{Spec} X_3 = \{2, 5, 7\}$$

$$v_1 = (0.6667, 0.3333, 0.6667) \in \text{Ker}(X_1 - 3I) \cap \text{Ker}(X_2 - 4I) \cap \text{Ker}(X_3 - 7I)$$

$$v_2 = (-0.2294, -0.6882, -0.6882) \in \text{Ker}(X_1 - I) \cap \text{Ker}(X_2 - 3I) \cap \text{Ker}(X_3 - 5I)$$

$$v_3 = (0.7071, 0, 0.7071) \in \text{Ker}(X_1 - 2I) \cap \text{Ker}(X_2 - 5I) \cap \text{Ker}(X_3 - 2I)$$

Then there exist

$$P = \begin{pmatrix} 4.5005 & 3.0003 & -4.5005 \\ 2.1796 & 0.0000 & -2.1796 \\ -2.1220 & -2.8289 & 3.5362 \end{pmatrix} = \begin{pmatrix} 0.6667 & -0.2294 & 0.7071 \\ 0.3333 & -0.6882 & 0.0000 \\ 0.6667 & -0.6882 & 0.7071 \end{pmatrix}^{-1}$$

such that

$$PX_1 P^{-1} = \begin{pmatrix} 3 & & \\ & 1 & \\ & & 2 \end{pmatrix}, PX_2 P^{-1} = \begin{pmatrix} 4 & & \\ & 3 & \\ & & 5 \end{pmatrix}, PX_3 P^{-1} = \begin{pmatrix} 7 & & \\ & 5 & \\ & & 2 \end{pmatrix}$$

(Calculations made with MatlabR2012b).

In this case, all possible matrices P diagonalizing X_i for some $i = 1, \dots, m$, (that they are such that $P = Q^{-1}$ where Q is a matrix whose columns are the eigenvectors corresponding to each of the eigenvalues of X_i for some $i = 1, \dots, n$), are matrices that diagonalizes all matrices simultaneously obtaining D_i or permutations of this matrices. In fact, we have the following proposition.

Proposition 2.14. If the set of matrices $\{X_1, \dots, X_m\}$ are simultaneously diagonalizable and for some i , X_i has simple eigenvalues, all matrices P diagonalizing X_i diagonalize X_j for all $j = 1 = \dots, m$.

Remark 2.15. If no matrix has simple eigenvalues then the result fails

Example 2.16. Let $T = (X_1, X_2, X_3)$ be a triple with

$$X_1 = \begin{pmatrix} 3 & 1 & -2 & -1 & 0 \\ 5 & -1 & -6 & 1 & 2 \\ 0 & 1 & 1 & -1 & 0 \\ 5 & -3 & -6 & 3 & 2 \\ 1 & -1 & -2 & 0 & 3 \end{pmatrix}, X_2 = \begin{pmatrix} 7 & 3 & -4 & -3 & 0 \\ 13 & -3 & -16 & 3 & 6 \\ 0 & 3 & 3 & -3 & 0 \\ 13 & -9 & -16 & 9 & 6 \\ 5 & -3 & -8 & 0 & 9 \end{pmatrix}, X_3 = \begin{pmatrix} 6 & 4 & -4 & -4 & 0 \\ 16 & -6 & -20 & 4 & 8 \\ 0 & 4 & 2 & -4 & 0 \\ 16 & -12 & -20 & 10 & 8 \\ 8 & -4 & -12 & 0 & 10 \end{pmatrix}$$

This triple diagonalize simultaneously, because there exists $Q \in \text{Gl}(n; \mathbb{C})$ with $Q = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 1 & -1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 0 & 1 & 1 & -1 \end{pmatrix}$,

$$\text{such that } P = Q^{-1} = \begin{pmatrix} -2 & 1 & 3 & 0 & -1 \\ 2 & -2 & -2 & 1 & 1 \\ 3 & -2 & -4 & 1 & 2 \\ 0 & 1 & 0 & -1 & 0 \\ 1 & 0 & -1 & 0 & 0 \end{pmatrix}$$

$$D_i = PX_iP^{-1}$$

$$D_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}, D_2 = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{pmatrix}, D_3 = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 6 \end{pmatrix}$$

But, in this case not all matrices diagonalizing one of this matrices diagonalize all set of matrices of m -tuple of simultaneously diagonalizable

matrices, because taking $P = \begin{pmatrix} -2 & 1 & 3 & 0 & -1 \\ 2 & -2 & -2 & 1 & 1 \\ 3 & -2 & -4 & 1 & 2 \\ -1 & 1 & 1 & -1 & 0 \\ 1 & 0 & -1 & 0 & 0 \end{pmatrix}$

$$\text{Then } P^{-1}X_1P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}, P^{-1}X_2P = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 6 & -1 \\ 0 & 0 & 0 & 0 & 7 \end{pmatrix}, P^{-1}X_3P = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 6 \end{pmatrix}$$

We observe that the matrix P only diagonalise X_3

3. Approximately simultaneously diagonalizable m -tuples of matrices

It is well known that near of a square matrix there is a diagonalizable matrix having simple eigenvalues. We ask if this result can be extended to the case of m -tuples of square matrices. We will try to obtain an answer using geometrical tools.

3.1. Group Lie action

The equivalence relation defined in (2.1) can be seen as the action over \mathfrak{M} in the following manner
Let us consider the following map

$$\alpha : Gl(n; \mathbb{C}) \times \mathfrak{M} \rightarrow \mathfrak{M} \\ (P, T) \rightarrow PTP^{-1} = (PX_1P^{-1}, \dots, PX_mP^{-1})$$

that verifies

- i) If $I \in Gl(n; \mathbb{C})$ is the identity element, then $\alpha(I, T) = T$ for all $T \in \mathfrak{M}$.
- ii) If P_1 and P_2 are in $Gl(n; \mathbb{C})$, then $\alpha(P_1, \alpha(P_2, T)) = \alpha(P_1P_2, T)$ for all $T \in \mathfrak{M}$.
 $\alpha(P_1, \alpha(P_2, T)) = \alpha(P_1, P_2TP_2^{-1}) = P_1P_2TP_2^{-1}P_1^{-1} = (P_1P_2)T(P_1P_2)^{-1} = \alpha(P_1P_2, T)$

So, the map α defines an action of $Gl(n; \mathbb{C})$ over \mathfrak{M} .

Fixing $T \in \mathfrak{M}$ we can consider the map

$$\alpha_T : Gl(n; \mathbb{C}) \rightarrow \mathfrak{M} \\ P \rightarrow \alpha_T(P) = \alpha(P, T)$$

We consider the following sets

$$\text{Im } \alpha_T = \mathcal{O}(T) = \{(Y_1, \dots, Y_m) = (PX_1P^{-1}, \dots, PX_mP^{-1}), \forall P \in Gl(n; \mathbb{C})\} \\ \text{Stab}(T) = \{P \in Gl(n; \mathbb{C}) \mid \alpha_T(P) = T\}$$

Fixing $P \in Gl(n; \mathbb{C})$ we can consider the map

$$\alpha_P : \mathfrak{M} \rightarrow \mathfrak{M} \\ T \rightarrow \alpha_P(T) = \alpha(P, T)$$

Notice that α_P is a bijection: if $\alpha(P, T_1) = \alpha(P, T_2)$ then $PT_1P^{-1} = PT_2P^{-1}$ and $T_1 = T_2$, so it is injective; for all $T \in \mathfrak{M}$, there exists $\bar{T} = P^{-1}TP$ such that $\alpha(P, \bar{T}) = T$, then it is surjective.

3.2. Approximately simultaneously diagonalizability

It is well known that close to any matrix there is a nearby that diagonalizes. Then the question is: given an m -tuple of square matrices, it is possible to found an m -tuple diagonalizing simultaneously?

In the case where that it is possible we say that the m -tuple is approximately simultaneously diagonalizable (abbreviated ASD), more concretely

Definition 3.1. [8] The m -tuple $T = (X_1, \dots, X_m)$ is approximately simultaneously diagonalizable if and only if for any $\epsilon > 0$, there exist a m -tuple of matrices (Y_1, \dots, Y_m) which are simultaneously diagonalizable and satisfy $\|Y_i - X_i\| < \epsilon$ for $i = 1, \dots, m$.

O'Meara and Vinsonhaler in [8], analyze approximately simultaneously diagonalizable matrices for the case where the matrices of the m -tuple commute.

Proposition 3.2. Let $T = (X_1, \dots, X_m)$ be an m -tuple simultaneously diagonalizable. Then, each $T^l \in \mathcal{O}(T)$ is an m -tuple simultaneously diagonalizable.

Proof. Taking into account that $T = (X_1, \dots, X_n)$ is an m -tuple simultaneously diagonalizable there exist $P \in Gl(n; \mathbb{C})$ such that $PTP^{-1} = (PX_1P^{-1}, \dots, PX_mP^{-1}) = (D_1, \dots, D_m) = D$ with D_i diagonal matrices for all $i = 1, \dots, m$.

Let $T' \in \mathcal{O}(T)$, then, there exist $P' \in Gl(n; \mathbb{C})$ such that $T' = P'T(P')^{-1} = (P'X_1(P')^{-1}, \dots, P'X_m(P')^{-1})$.

So, $T' = P'P^{-1}DP(P')^{-1} = (P'P^{-1}D_1P(P')^{-1}, \dots, P'P^{-1}D_mP(P')^{-1}) = (P''D_1(P'')^{-1}, \dots, P''D_m(P'')^{-1})$, with $P'' = P'P^{-1} \in Gl(n; \mathbb{C})$. \square

Consequently, and taking into account that if $T' \in \mathcal{O}(T)$ is $\mathcal{O}(T) = \mathcal{O}(T')$, we can use miniversal deformations to study approximately simultaneously diagonalizability.

3.3. Miniversal deformations

Definition 3.3. A deformation of an element $X_0 \in \mathfrak{M}$ is a family of elements of \mathfrak{M} indexed by $\lambda \in \Lambda$ $\varphi : \Lambda \rightarrow \mathfrak{M}$ where $\Lambda \subset \mathbb{F}^\ell$ is a neighborhood of 0, and where $\varphi(0) = X_0$ and φ depends holomorphically (smoothly) on the parameters.

Definition 3.4. A deformation $\varphi(\lambda) = \varphi(\lambda_1, \dots, \lambda_\ell)$ of X_0 is versal if and only if for any deformation $\varphi'(\mu_1, \dots, \mu_k) \in \mathfrak{M}$ of X_0 , $\varphi'(\mu)$ is induced by $\varphi(\lambda)$, i.e., there exists a neighborhood V of 0 in \mathbb{F}^k , a map $\psi : V \rightarrow \mathbb{F}^\ell$ with $\psi(0) = 0$, and a map $g : V \rightarrow G$ with $g(0) = I$ such that $\forall \mu \in V$, $\varphi'(\mu) = g(\mu)\varphi(\psi(\mu))g^{-1}(\mu)$ with ψ and g holomorphic (smooth).

It is obvious that if we have a versal deformation of an element automatically we have a versal deformation of any element that is equivalent to it, since if $X = \alpha(g, X_0)$ is an equivalent element of X_0 and $\varphi(\lambda)$ is a versal deformation of X' then $\alpha(g^{-1}, X(\lambda))$ is a versal deformation of X_0 .

A versal deformation having minimal number of parameters is called *miniversal*.

The following result was proved by Arnold [10], in the case where $Gl(n; \mathbb{C})$ acts on $M_n(\mathbb{C})$, and was generalized by Tannenbaum [11], in the case where a Lie group acts on a complex manifold. It provides the relationship between a versal deformation of X_0 and the local structure of the orbit.

Theorem 3.5 ([11]). 1. A deformation $\varphi(\lambda)$ of (X_0) is versal if and only if it is transversal to the orbit $\mathcal{O}(X_0)$ at (X_0) .

2. Minimal number of parameters of a versal deformation is equal to the codimension of the orbit of X_0 in \mathfrak{M} , $\ell = \text{codim } \mathcal{O}(X_0)$.

Corollary 3.6. Then $\varphi(\lambda) = X_0 + (T_{X_0}\mathcal{O}(X_0))^\perp$ for some scalar product is a miniversal deformation.

Let $d\alpha_{X_0} : T_T\mathcal{G} \rightarrow \mathfrak{M}$ be the differential of α_{X_0} at the unit element I . It is easy to compute $d\alpha_{X_0}(P)$:

$$d\alpha_T(P) = ([X_1, P], \dots, [X_m, P]) \in \mathfrak{M}, \quad P \in T_T\mathcal{G}.$$

If we define scalar products in \mathfrak{M} and $T_T\mathcal{G}$, we can consider the adjoint application of $d\alpha_{X_0}$. The Euclidean scalar products considered in this paper are defined as follows:

For all $T_i = (X_1^i, \dots, X_m^i) \in \mathfrak{M}$ and for all $P_i \in T_T\mathcal{G}$

$$\begin{aligned} \langle T_1, T_2 \rangle_1 &= \text{trace}(X_1^1 X_1^{2*}) + \dots + \text{trace}(X_1^m X_1^{m*}), \\ \langle P_1, P_2 \rangle_2 &= \text{trace}(P_1 P_2^*), \end{aligned}$$

where X^* denotes the conjugate transpose of a matrix X .

The adjoint linear mapping $d\alpha_T^* : \mathfrak{M} \rightarrow T_T\mathcal{G}$ is defined by the relation

$$\langle d\alpha_{X_0}(P), Z \rangle_1 = \langle P, d\alpha_{X_0}^*(Z) \rangle_2, \quad P \in T_T\mathcal{G}, \quad Z \in \mathfrak{M}.$$

It is straightforward to find

$$d\alpha_{X_0}^*(W) = ([X^*, A_0] + [Y^*, B_0] + [Z^*, C_0]) \in T_T\mathcal{G}, \quad W = (X, Y, Z) \in \mathfrak{M}.$$

The mappings $d\alpha_{X_0}$ and $d\alpha_{X_0}^*$ provide a simple description of the tangent spaces $T_{X_0}\mathcal{O}(X_0)$, $T_T\mathcal{S}tab(X_0)$ and their normal complements $(T_{X_0}\mathcal{O}(X_0))^\perp$, $(T_T\mathcal{S}(X_0))^\perp$.

Theorem 3.7. The tangent spaces to the orbit of the m -tuple of matrices T and the corresponding normal complementary subspace can be found in the following form

1. $T_T\mathcal{O}(X_0) = \text{Im } d\alpha_T \subset \mathfrak{M}$.
2. $(T_T\mathcal{O}(X_0))^\perp = \text{Ker } d\alpha_T^* \subset \mathfrak{M}$.

After this theorem, it is easy to compute these spaces.

Corollary 3.8. 1. $T_{X_0}\mathcal{O}(X_0) = \{([P, A_0], [P, B_0], [P, C_0]) \mid P \in T_T\mathcal{G}\}$

2. $(T_{X_0}\mathcal{O}(X_0))^\perp = \{(X, Y, Z) \in \mathfrak{M} \mid [X^*, A_0] + [Y^*, B_0] + [Z^*, C_0] = 0\}$

Remark 3.9. Let $X_0 = (X_1^0, \dots, X_m^0)$ be an n -tuple of matrices and we consider $\mathfrak{X}_i = (0, \dots, 0, X_i, 0, \dots, 0)$ an m -tuple of matrices such that $X_i^0 + X_i$ is a miniversal deformation of X_i^0 . Then $\mathfrak{X}_i \in (T_{X_0}\mathcal{O}(X_0))^\perp$ and consequently $\mathfrak{X} = \sum \mathfrak{X}_i \in T_{X_0}\mathcal{O}(X_0)^\perp$.

Consequently, we have the following proposition.

Proposition 3.10. Let $X_0 = (X_1^0, \dots, X_m^0)$ be an n -tuple of matrices. Then, for all $\varepsilon > 0$ there exist $\mathfrak{X} = (X_1, \dots, X_m)$ such that $X_i^0 + X_i$ is diagonalizable, for all $i = 1, \dots, m$.

Remark 3.11. Given any n -tuple of matrices, we can find in a neighborhood, an n -tuple of matrices in which all matrices are diagonalizable but not necessarily all matrices in the n -tuple diagonalize simultaneously.

Example 3.12. Consider the following pair of matrices $\left(\begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}\right)$ that they are no diagonalizable, and the following family of perturbations of the pair: $\left(\begin{pmatrix} 2+\varepsilon_1 & 0 \\ 1 & 2+\varepsilon_2 \end{pmatrix}, \begin{pmatrix} 3+\varepsilon_3 & 1 \\ 0 & 3+\varepsilon_4 \end{pmatrix}\right)$ for all ε_i with $\varepsilon_1 \neq \varepsilon_2$ and $\varepsilon_3 \neq \varepsilon_4$.

Clearly, both matrices are diagonalizable.

For simultaneously diagonalization it is necessary that both matrices commute, but

$$\begin{pmatrix} 2+\varepsilon_1 & 0 \\ 1 & 2+\varepsilon_2 \end{pmatrix} \begin{pmatrix} 3+\varepsilon_3 & 1 \\ 0 & 3+\varepsilon_4 \end{pmatrix} \neq \begin{pmatrix} 3+\varepsilon_3 & 1 \\ 0 & 3+\varepsilon_4 \end{pmatrix} \begin{pmatrix} 2+\varepsilon_1 & 0 \\ 1 & 2+\varepsilon_2 \end{pmatrix}$$

for all $\varepsilon_1, \varepsilon_2$, so both matrices diagonalize but not diagonalize simultaneously.

Now, we consider the following perturbation $\left(\begin{pmatrix} 2 & \varepsilon_1 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ \varepsilon_2 & 3 \end{pmatrix}\right)$ for all ε_i with $\varepsilon_1 \cdot \varepsilon_2 \neq 0$. Clearly, both matrices are diagonalizable.

Analyzing commutativity

$$\begin{pmatrix} 2 & \varepsilon_1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ \varepsilon_2 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ \varepsilon_2 & 3 \end{pmatrix} \begin{pmatrix} 2 & \varepsilon_1 \\ 1 & 2 \end{pmatrix}$$

equivalently $\varepsilon_1 \varepsilon_2 = 1$

So, taking $\varepsilon_1 = \varepsilon_2 = 1$, both matrices diagonalize simultaneously, (it suffices to consider $P^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$)

The near pair of matrices in this family diagonalizing simultaneously is with $(\varepsilon_1, \varepsilon_2)$ minimizing distance of the variety $V = \{(\varepsilon_1, \varepsilon_2) \mid \varepsilon_1 \varepsilon_2 = 1\}$

In general, a lower bound at the distance of the a n -tuple of matrices to a one n -tuple diagonalizing simultaneously is given in the following proposition

Proposition 3.13. Let $T = (X_1, \dots, X_m)$ be a n -tuple of matrices and $T(\lambda) = (X_1(\lambda), \dots, X_m(\lambda))$ with $\lambda \in \mathbb{C}^\ell$ a family of n -tuples such that in a neighborhood of $0 \in \mathbb{C}^\ell$ is a miniversal deformation of the given n -tuple. A lower bound at the distance of the a n -tuple of matrices to a one n -tuple diagonalizing simultaneously is

$$\inf\{\text{dist}(0, \lambda), 0, \lambda \in \mathbb{C}^\ell \mid X_i(\lambda)X_j(\lambda) = X_j(\lambda)X_i(\lambda) \forall 1 \leq i, j \leq m\}$$

Example 3.14. Let $T = \left(\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right)$ be a triple of matrices no diagonalizing. Let us consider the family of triples $T(\varepsilon) = \left(\begin{pmatrix} \varepsilon_1 & \varepsilon_3 \\ 1+\varepsilon_2 & \varepsilon_4 \end{pmatrix}, \begin{pmatrix} \varepsilon_5 & 1 \\ \varepsilon_6 & \varepsilon_7 \end{pmatrix}, \begin{pmatrix} \varepsilon_8 & 0 \\ 0 & 1+\varepsilon_9 \end{pmatrix}\right)$ in such a way that for some ε with $\|\varepsilon\| > 0$ is a miniversal (no orthogonal) deformation.

The subset of the commuting triples in the family is

$$V = \left\{ T \in \{T \in T(\varepsilon) \mid \begin{aligned} \varphi_1(\varepsilon) &= \varepsilon_3 \cdot \varepsilon_6 - \varepsilon_2 - 1 &= 0, \\ \varphi_2(\varepsilon) &= \varepsilon_1 - \varepsilon_4 - \varepsilon_3 \cdot \varepsilon_5 + \varepsilon_3 \cdot \varepsilon_7 &= 0, \\ \varphi_3(\varepsilon) &= \varepsilon_4 \cdot \varepsilon_6 - \varepsilon_1 \cdot \varepsilon_6 + \varepsilon_5 \cdot (\varepsilon_2 + 1) - \varepsilon_7 \cdot (\varepsilon_2 + 1) &= 0, \\ \varphi_4(\varepsilon) &= \varepsilon_9 - \varepsilon_8 + 1 &= 0. \end{aligned} \right\}$$

We can compute the minimal distance by means the Lagrange's undetermined multipliers method, from the function:

$$f(\varepsilon, \lambda) = \sum_{i=1}^9 \varepsilon_i^2 + \sum_{i=1}^4 \lambda_i \varphi_i(\varepsilon)$$

The minimal distance is $\sqrt{3/2}$, a triple minimizing this distance is a triple of commuting matrices with $\varepsilon_2 = -1, \varepsilon_8 = 1/2 = -\varepsilon_9$ and $\varepsilon_i = 0$ for $i = 1, 3, 4, 5, 6, 7$ but no diagonalize simultaneously.

Taking the solution $\varepsilon_3 = \varepsilon_6 = \sqrt{2}, \varepsilon_2 = 1, \varepsilon_8 = \frac{1}{2} = -\varepsilon_9$, and $\varepsilon_i = 0$ for $i = 1, 4, 5, 7$ with distance $\sqrt{11/2}$ we have a triple of commuting matrices and they diagonalize simultaneously with $P^{-1} = \begin{pmatrix} 1 & 1 \\ 2^{1/4} & -2^{1/4} \end{pmatrix}$.

References

- [1] M. I. Garcia-Planas, M.D. Magret, *Polynomial Matrices and Generalized Linear Multivariable Dynamical Systems*, N. Mastorakis (editor) *Recent Advances in Applied and Theoretical Mathematics*, Wseas Press, Athens, 2000, pp. 17-22.
- [2] M. I. Garcia-Planas, S. Tarragona, *Perturbation Analysis of Eigenvalues of Polynomial Matrices Smoothly Depending on Parameters*, N. Mastorakis, V. Mladenov et Al. (Eds), *Recent Researches in System Science*, Wseas Press, Athens, 2011, pp. 100-105.
- [3] J. C. Zúñiga-Anaya, *Structural properties of polynomial and rational matrices, a survey*, Math. AETerna, **1**(06) (2011), 361-403.
- [4] R. Guralnick, *A note on commuting pairs of matrices*, Linear Multilinear Algebra, **31** (1992), 71-75.
- [5] Y. Han, *Commuting triples of matrices*. Electron. J. Linear Algebra, **13** (2005), 274-343.
- [6] A. Marrani, P. Truini, *Exceptional Lie algebras, SU(3) and Jordan pairs part 2: Zorn-type representations*, (2014), arXiv:1403.5120v2.
- [7] S. Okubo, *Introduction to Octonion and Other Non-Associative Algebras in Physics*, Cambridge University Press, (1995).
- [8] K. C. O'Meara, C. Vinsonhaler, *On approximately simultaneously diagonalizable matrices*, Linear Algebra Appl., **412**(1) (2006), 39-74.
- [9] Sh. Friedland, *Simultaneous similarity of matrices*, Adv. Math., **50** (1983), 189-265.
- [10] V. I. Arnold, *On matrices depending on parameters*, Russian Math. Surveys, **26**(2) (1971), 29-43.
- [11] A. Tannenbaum, *Invariance and System Theory: Algebraic and geometric Aspects*, Lect. Notes in Math. **845**, Springer-Verlag, (1981).