Similarity of Families of Matrices
An article in CRC Handbook of Linear Algebra

Shmuel Friedland

Department of Mathematics, Statistics, and Computer Science,

University of Illinois at Chicago

Chicago, Illinois 60607-7045, USA

November 15, 2005

In this article we use the notations, definitions and the facts given in our article "Matrices over Integral Domains". The aim of this article to introduce the reader with the two main difficult problems in matrix theory:

1. Similarity of matrices over integral domains which are not fields.

2. Simultaneous similarity of tuples of matrices over $\mathbb{C}$.

Problem *1* is notoriously difficult. We show that for the local ring H_0 this problem reduces to a Problem *2* for certain kind of matrices. We then discuss certain special cases of Problem *2* as simultaneous similarity of tuples of matrices to upper triangular and diagonal matrices. L-property of pairs of matrices, that discussed next, is closely related to simultaneous similarity of pair of matrices to a diagonal pair. The rest of the article is devoted to a "solution" of the Problem *2*, by the author, in terms of basic notions of algebraic geometry.

1 Similarity of matrices

The classical result of Weierstrass [Wei67] claims that the similarity class of $A \in F^{n\times n}$ is determined by the the invariant factors of $-A + xI_n$ over $F[x]$. (See §? .) For a given $A, B \in F^{n\times n}$ one can easily determine if A and B are similar, by considering *only* the ranks of certain 3 associated matrices with A, B [GB77]. It is well known that it is a difficult problem to determine if $A, B \in \mathbb{D}^{n\times n}$ are $\mathbb{D}$-similar for most of integral domains which are not fields. Then emphasize of this section is the similarity over the local field H_0. The subject of similarity of matrices over H_0 arises naturally in theory linear differential equations having singularity with respect to a parameter. It was studied by Wasow in [Was63], [Was77] and [Was78].

The standard results on the tensor products can be found in [MM64]. Most of the results of this section related to the analytic similarity over H_0 are taken from [Fri80].

Definitions:

For $E = (e_{ij}) \in \mathbb{D}^{m\times n}, G \in \mathbb{D}^{p\times q}$ denote by $E \otimes G \in \mathbb{D}^{mp\times nq}$, the **tensor** or **Kronecker** product of E with G. It can be viewed as $m \times n$ block matrix $[e_{ij}G]_{i,j=1}^{m,n}$. Alternatively $E \otimes G$ can be identified with the linear operator

$$L(E, G) : \mathbb{D}^{q\times n} \to \mathbb{D}^{p\times m}, \quad X \to GXE^{\mathrm{T}}.$$

$A, B \in \mathbb{D}^{m\times m}$ are called **similar**, denoted by $A \approx B$ if $B = QAQ^{-1}$, for some $Q \in \mathbf{GL}_m(\mathbb{D})$.

Let $A, B \in \mathrm{H}(\Omega)^{n\times n}$. Then A and B are called **analytically similar**, denoted as $A \overset{a}{\approx} B$, if A and B are similar over $\mathrm{H}(\Omega)$.

A and B are called **locally similar** if for any $\zeta \in \Omega$, the restrictions A_ζ, B_ζ of A, B to the local rings H_ζ respectively are similar over H_ζ.

A, B are called **point-wise** similar if $A(\zeta), B(\zeta)$ are similar matrices in $\mathbb{C}^{n\times n}$ for each $\zeta \in \Omega$.

A, B are called **rationally** similar, denoted as $A \overset{r}{\approx} B$, if A, B are similar over the quotient field $\mathcal{M}(\Omega)$ of $\mathrm{H}(\Omega)$.

Let $A, B \in \mathrm{H}_0^{n\times n}$:

$$A(x) = \sum_{k=0}^{\infty} A_k x^k, \ |x| < R(A), \quad B(x) = \sum_{k=0}^{\infty} B_k x^k, \ |x| < R(B).$$

Then $\eta(A, B)$ and $\mathcal{K}_p(A, B)$ are the index and the number of local invariant polynomials of degree p of the matrix $I_n \otimes A(x) - B(x)^{\mathrm{T}} \otimes I_n$ respectively, for $p = 0, 1, \ldots$.

$\lambda(x)$ is called an **algebraic** function if there exists a monic polynomial $p(\lambda, x) = \lambda^n + \sum_{i=1}^{n} q_i(x)\lambda^{n-i} \in (\mathbb{C}[x])[\lambda]$ of λ-degree $n \geq 1$ such that $p(\lambda(x), x) = 0$ identically. Then $\lambda(x)$ is a multi-valued function on $\mathbb{C}$ which has n **branches**. At each point $\zeta \in \mathbb{C}$ each branch $\lambda_j(x)$ of $\lambda(x)$ has **Puiseaux** expansion: $\lambda_j(x) = \sum_{i=0}^{\infty} b_{ji}(\zeta)(x-\zeta)^{\frac{i}{m}}$, which converges for $|x - \zeta| < R(\zeta)$, and some integer m depending on $p(x)$. $\sum_{i=0}^{m} b_{ji}(\zeta)(x-\zeta)^{\frac{i}{m}}$ is called the **linear part** of $\lambda_j(x)$ at ζ. Two distinct branches $\lambda_j(x)$ and $\lambda_k(x)$ are called **tangent** at $\zeta \in \mathbb{C}$ if the linear parts of $\lambda_j(x)$ and $\lambda_k(x)$ coincide at ζ. Each branch $\lambda_j(x)$ has Puiseaux expansion around ∞: $\lambda_j(x) = x^l \sum_{i=0}^{\infty} c_{ji} x^{-\frac{i}{m}}$, which converges for $|x| > R$. Here l is the smallest nonnegative integer such that $c_{j0} \neq 0$ at least for some branch λ_j. $x^l \sum_{i=0}^{m} c_{ji} x^{-\frac{i}{m}}$ is called the **principal part** of $\lambda_j(x)$ at ∞. Two distinct branches $\lambda_j(x)$ and $\lambda_k(x)$ are called **tangent** at ∞ if the principal parts of $\lambda_j(x)$ and $\lambda_k(x)$ coincide at ∞.

Facts:

1. The similarity relation is an equivalence relation on $\mathbb{D}^{m\times m}$.

2. $A \approx B \iff A(x) = -A + xI \overset{s}{\sim} B(x) = -B + xI$.

3. Let $A, B \in F^{n\times n}$. Then A and B are similar if and only if the pencils $-A + xI$ and $-B + xI$ have the same invariant polynomials over $F[x]$.

$(E \otimes G)(U \otimes V) = EU \otimes GV$ whenever the products EU and GV are defined. Also $(E \otimes G)^{\mathrm{T}} = E^{\mathrm{T}} \otimes G^{\mathrm{T}}$.

4. For $A, B \in \mathbb{D}^{n\times n}$

$$A \approx B \Rightarrow \tag{1.1}$$
$$I_n \otimes A - A^{\mathrm{T}} \otimes I_n \sim I_n \otimes A - B^{\mathrm{T}} \otimes I_n \sim I_n \otimes B - B^{\mathrm{T}} \otimes I_n.$$

5. [Gur80]: There are examples over Euclidean domains for which the reverse implications in (1.1) does not hold.

6. [GB77]: For $\mathbb{D} = F$ the reverse implication in (1.1) holds .

7. [Fri80]: Let $A \in F^{m\times m},\ B \in F^{n\times n}$. Then

$$\text{null } (I_n \otimes A - B^{\mathrm{T}} \otimes I_m) \le \frac{1}{2}(\text{null } (I_m \otimes A - A^{\mathrm{T}} \otimes I_m) + \text{null } (I_n \otimes B - B^{\mathrm{T}} \otimes I_n)).$$

   Equality holds if and only if $m = n$ and A and B are similar.

8. Let $A \in F^{n\times n}$ and assume that $p_1(x), \ldots, p_k(x) \in F[x]$ are the nontrivial normalized invariant polynomials of $-A + xI$, where $p_1(x)|p_2(x)|\ldots|p_k(x)$ and $p_1(x)p_2(x)\ldots p_k(x) = \det\,(xI - A)$. Then $A \approx C(p_1) \oplus C(p_2) \oplus \ldots \oplus C(p_k)$ and $C(p_1) \oplus C(p_2) \oplus \ldots \oplus C(p_k)$ is called the **rational canonical form** of A.

9. For $A, B \in \mathrm{H}(\Omega)^{n\times n}$ the analytic similarity implies the local similarity, the local similarity implies implies the point-wise similarity, and the point-wise similarity implies rational similarity.

10. For $n = 1$ the all the four concepts are equivalent. For $n \ge 2$ local similarity, point-wise similarity and rational similarity are distinct.

11. The equivalence of the three matrices in (1.1) over $\mathrm{H}(\Omega)$ implies the point-wise similarity of A and B.

12. Let $A, B \in \mathrm{H}_0^{n\times n}$. Then A and B are analytically similar over H_0 if and only if A and B are rationally similar over H_0 and there exists $\eta(A,A)+1$ matrices $T_0, ..., T_\eta \in \mathbb{C}^{n\times n}$ $(\eta = \eta(A,A))$, such that det $T_0 \neq 0$ and

$$\sum_{i=0}^{k} A_i T_{k-i} - T_{k-i} B_i = 0, \quad k = 0, ..., \eta(A,A). \tag{1.2}$$

13. Suppose that the characteristic polynomial of $A(x)$ splits over H_0:

$$\det\,(\lambda I - A(x)) = \prod_{i=1}^{n} (\lambda - \lambda_i(x)), \quad \lambda_i(x) \in \mathrm{H}_0,\ i = 1, ..., n. \tag{1.3}$$

Then $A(x)$ is analytically similar to

$$C(x) = \oplus_{i=1}^{\ell} C_i(x), \quad C_i(x) \in \mathrm{H}_0^{n_i\times n_i},$$
$$(\alpha_i I_{n_i} - C_i(0))^{n_i} = 0,\ \alpha_i = \lambda_{n_i}(0),\ \alpha_i \neq \alpha_j \text{ for } i \neq j,\ i,j = 1, ..., \ell. \tag{1.4}$$

14. Assume that the characteristic polynomial of $A(x) \in \mathrm{H}_0$ splits in H_0.Then $A(x)$ is analytically similar to a block diagonal matrix $C(x)$ of the form (1.4) such that each $C_i(x)$ is an upper triangular matrix whose off-diagonal entries are polynomial in x. Moreover, the degree of each polynomial entry above the diagonal in the matrix $C_i(x)$ does not exceed $\eta(C_i, C_i)$ for $i = 1, ..., \ell$.

15. Let $P(x)$ and $Q(x)$ be matrices of the form (1.4)

$$P(x) = \oplus_{i=1}^{p} P_i(x),\ P_i(x) \in \mathrm{H}_0^{m_i\times m_i},$$
$$(\alpha_i I_{m_i} - P_i(0))^{m_i} = 0,\ \alpha_i \neq \alpha_j \text{ for } i \neq j,\ i,j = 1, ..., p,$$
$$Q(x) = \oplus_{j=1}^{q} Q_j(x),\ Q_j(x) \in \mathrm{H}_0^{n_j\times n_j},$$
$$(\beta_j I_{n_j} - Q_j(0))^{n_j} = 0,\ \beta_i \neq \beta_j \text{ for } i \neq j,\ i,j = 1, ..., q. \tag{1.5}$$

Assume furthermore that

$$\alpha_i = \beta_i,\ i = 1, ..., t,\ \alpha_j \neq \beta_j,\ i = t{+}1, ..., p,\ j = t{+}1, ..., q,\ 0 \leq t \leq \min(p, q). \tag{1.6}$$

Then the nonconstant local invariant polynomials of $I \otimes P(x) - Q(x)^{\mathrm{T}} \otimes I$ are the nonconstant local invariant polynomials of $I \otimes P_i(x) - Q_i(x)^{\mathrm{T}} \otimes I$ for $i = 1, ..., t$:

$$\mathcal{K}_p(P, Q) = \sum_{i=1}^{t} \mathcal{K}_p(P_i, Q_i), \quad p = 1, ..., . \tag{1.7}$$

In particular if $C(x)$ is of the form (1.4) then

$$\eta(C, C) = \max_{1 \leq i \leq \ell} \eta(C_i, C_i). \tag{1.8}$$

16. $A(x) \overset{a}{\approx} B(x) \iff A(y^m) \overset{a}{\approx} B(y^m)$ for any $2 \leq m \in \mathbb{N}$.

17. **Weierstrass preparation theorem** [GR65]: For any monic polynomial $p(\lambda, x) = \lambda^n + \sum_{i=1}^{n} a_i(x)\lambda^{n-i} \in \mathrm{H}_0[\lambda]$ there exists $m \in \mathbb{N}$ such that $p(\lambda, y^m)$ splits over H_0.

18. For a given rational canonical form $A(x) \in \mathrm{H}_0^{2\times 2}$ there are at most countable number of analytic similarity classes. (See Example *3*.)

19. For given rational canonical form $A(x) \in \mathrm{H}_0^{n\times n}$, where $n \geq 3$, there may exists a family of distinct similarity classes correspond to a finite dimensional variety. (See Example *4*.)

20. Let $A(x) \in H_0^{n\times n}$ and assume that the characteristic polynomial of $A(x)$ splits in H_0 as in (1.3). Let $B(x) = \mathrm{diag}\,(\lambda_1(x), ..., \lambda_n(x))$. Then $A(x)$ and $B(x)$ are not analytically similar if and only if there exists a nonnegative integer p such that

$$\begin{aligned} &\mathcal{K}_p(A, A) + \mathcal{K}_p(B, B) < 2\mathcal{K}_p(A, B), \\ &\mathcal{K}_j(A, A) + \mathcal{K}_j(B, B) = 2\mathcal{K}_j(A, B),\ j = 0, ..., p-1, \quad \text{if } p \geq 1. \end{aligned} \tag{1.9}$$

In particular $A(x)\overset{a}{\approx}B(x)$ if and only if the three matrices given in (1.1) are equivalent over H_0.

21. [Fri78]: Let $A(x) \in \mathbb{C}[x]^{n\times n}$. Then each eigenvalue $\lambda(x)$ of $A(x)$ is an algebraic function. Assume that $A(\zeta)$ is diagonable for some $\zeta \in \mathbb{C}$. Then the linear part of each branch of $\lambda_j(x)$ is linear at ζ, i.e. is of the form $\alpha + \beta x$ for some $\alpha, \beta \in \mathbb{C}$.

22. Let $A(x) \in \mathbb{C}[x]^{n\times n}$ be of the form $A(x) = \sum_{k=0}^{\ell} A_k x^k$, where $A_k \in \mathbb{C}^{n\times n}$ for $k = 0, ..., \ell$ and $\ell \geq 1,\ A_\ell \neq 0$. Then one of the following conditions imply that $A(x) = S(x)B(x)S^{-1}(x)$, where $S(x) \in \mathbf{GL}(n, \mathbb{C}[x])$ and $B(x) \in \mathbb{C}[x]^{n\times n}$ is a diagonal matrix of the form $\sum_{i=1}^{m} \lambda_i(x) I_{k_i}$, where $k_1, \ldots, k_m \geq 1$. Furthermore $\lambda_1(x), ..., \lambda_m(x)$ are m distinct polynomials satisfying the following conditions:

   (a) $\deg \lambda_1 = \ell \geq \deg \lambda_i(x),\ i = 2, ..., m-1.$

   (b) The polynomial $\lambda_i(x) - \lambda_j(x)$ has only simple roots in $\mathbb{C}$ for $i \neq j$. $(\lambda_i(\zeta) = \lambda_j(\zeta) \Rightarrow \lambda_i'(\zeta) \neq \lambda_j'(\zeta))$.

   I. The characteristic polynomial of $A(x)$ splits in $\mathbb{C}[x]$, i.e. all the eigenvalues of $A(x)$ are polynomials. $A(x)$ is point-wise diagonable in $\mathbb{C}$ and no two distinct eigenvalues are tangent at any $\zeta \in \mathbb{C}$.

   II. $A(x)$ is point-wise diagonable in $\mathbb{C}$ and A_ℓ is diagonable. No two distinct eigenvalues are tangent at any point $\zeta \in \mathbb{C} \cup \{\infty\}$. Then $A(x)$ is strictly similar to $B(x)$, i.e. $S(x)$ can be chosen in $\mathbf{GL}(n, \mathbb{C})$. Furthermore $\lambda_1(x), ..., \lambda_m(x)$ satisfy the additional condition:

   (c) For $i \neq j$ either $\frac{d^\ell \lambda_i}{d^\ell x}(0) \neq \frac{d^\ell \lambda_j}{d^\ell x}(0)$ or $\frac{d^\ell \lambda_i}{d^\ell x}(0) = \frac{d^\ell \lambda_j}{d^\ell x}(0)$ and $\frac{d^{\ell-1} \lambda_i}{d^{\ell=1} x}(0) \neq \frac{d^{\ell-1} \lambda_j}{d^{\ell-1} x}(0)$.

Examples:

1. Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}, \ B = \begin{bmatrix} 1 & 1 \\ 0 & 5 \end{bmatrix} \in \mathbb{Z}^{2\times 2}.$$

Then $A(x)$ and $B(x)$ have the same invariant polynomials over $\mathbb{Z}[x]$ and A and B are not similar over $\mathbb{Z}$.

2. Let

$$A(z) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad D(z) = \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix}.$$

Then $zA(z) = D(z)A(z)D(z)^{-1}$, i.e. $A(z), zA(z)$ are rationally similar. Clearly $A(z)$ and $zA(z)$ are not point-wise similar for any Ω containing 0. Now $zA(z), z^2A(z)$ are point-wise similar in $\mathbb{C}$, but they are not locally similar on H_0.

3. Let $A(x) \in \mathrm{H}_0^{2\times 2}$ and assume that $\det(\lambda I - A(x)) = (\lambda - \lambda_1(x))(\lambda - \lambda_2(x))$. Then $A(x)$ is analytically similar either to a diagonal matrix or to

$$B(x) = \begin{bmatrix} \lambda_1(x) & x^k \\ 0 & \lambda_2(x) \end{bmatrix}, \quad k = 0, ..., p \ (p \geq 0).$$

Furthermore if $A(x) \overset{a}{\approx} B(x)$ then $\eta(A, A) = k$.

4. Let $A(x) \in \mathrm{H}_0^{3\times 3}$. Assume that

$$A(x) \overset{r}{\approx} C(p), \quad p(\lambda, x) = \lambda(\lambda - x^{2m})(\lambda - x^{4m}), \quad m \geq 1.$$

Then $A(x)$ is analytically similar to a matrix

$$B(x, a) = \begin{bmatrix} 0 & x^{k_1} & a(x) \\ 0 & x^{2m} & x^{k_2} \\ 0 & 0 & x^{4m} \end{bmatrix}, \quad 0 \leq k_1, k_2 \leq \infty \ (x^\infty = 0),$$

where $a(x)$ is a polynomial of degree $4m-1$ at most. Furthermore $B(x,a)\overset{a}{\approx}B(x,b)$ if and only if

(1) if $a(0) \neq 1$ then $b-a$ is divisible by x^m.

(2) if $a(0)=1$ and $\frac{d^i a}{dx^i}=0,\ i=1,...,k-1,\ \frac{d^k a}{dx^k}\neq 0$ for $1 \le k < m$ then $b-a$ is divisible by x^{m+k}.

(3) if $a(0)=1$ and $\frac{d^i a}{dx^i}=0,\ i=1,...,m$ then $b-a$ is divisible by x^{2m}.

Then for $k_1=k_2=m$ and $a(0)\in\mathbb{C}\backslash\{1\}$ we can assume that $a(x)$ is a polynomial of degree less than m. Furthermore the similarity classes of $A(x)$ is uniquely determined by such $a(x)$. These similarity classes are parameterized by $\mathbb{C}\backslash\{1\}\times\mathbb{C}^{m-1}$ (the Taylor coefficients of $a(x)$).

2 Simultaneous similarity of matrices

In this section we introduce the notion of simultaneous similarity of matrices over $\mathbb{D}$. The problem of simultaneous similarity of matrices over a field $\mathbb{F}$, i.e. to describe the similarity class of a given m (≥ 2) tuple of matrices or to decide when a given two tuples of matrices are simultaneously similar, is in general a hard problem, which will be discussed in the next sections. There are some cases where this problem has a relatively simple solution. As shown below the problem of analytic similarity of $A(x), B(x) \in \mathrm{H}_0^{n\times n}$ reduces to the problem of simultaneously similarity of certain 2-tuples of matrices.

Definitions:

For $A_0,...,A_l\in\mathbb{D}^{n\times n}$ denote by $\mathcal{A}(A_0,...,A_l)\subset\mathbb{D}^{n\times n}$ the minimal algebra in $\mathbb{D}^{n\times n}$ containing I_n and $A_0,...,A_l$. Thus every matrix $G\in\mathcal{A}(A_0,...,A_l)$ is a noncommutative polynomial in $A_0,...,A_l$.

For $l\geq 1$, $(A_0,A_1,...,A_\ell),(B_0,...,B_\ell)\in(\mathbb{D}^{n\times n})^{\ell+1}$ are called **simultaneously** similar, denoted by $(A_0,A_1,...,A_\ell)\approx(B_0,...,B_\ell)$, if there exists $P\in\mathbf{GL}(n,\mathbb{D})$ such that $B_i=PA_iP^{-1}, i=0,...,\ell$, i.e. $(B_0,B_1,...,B_\ell)=P(A_0,A_1,...,A_\ell)P^{-1}$.

Associate with $(A_0, A_1, ..., A_\ell), (B_0, ..., B_\ell) \in (\mathbb{D}^{n\times n})^{\ell+1}$ matrix polynomials $A(x) = \sum_{i=0}^{\ell} A_i x^i, B(x) = \sum_{i=0}^{\ell} B_i x^i \in \mathbb{D}[x]^{n\times n}$. $A(x)$ and $B(x)$ are called **strictly** similar ($A \overset{s}{\approx} B$) if there exists $P \in \mathbf{GL}(n, \mathbb{D})$ such that $B(x) = PA(x)P^{-1}$.

Clearly $A \overset{s}{\approx} B \iff (A_0, A_1, ..., A_\ell) \approx (B_0, ..., B_\ell)$.

Facts:

1. Let $l \in \mathbb{N}$, $(A_0, \dots, A_l), (B_0, \dots, B_l) \in (\mathbb{C}^{n\times n})^{l+1}$, and $U = (U_{ij})_{i,j=1}^{l+1}, V = (V_{ij})_{i,j=1}^{l+1}, W = (W_{ij})_{i,j=1}^{l+1} \in \mathbb{C}^{n(l+1)\times n(l+1)}$, $U_{ij}, V_{ij}, W_{ij} \in \mathbb{C}^{n\times n}, i, j = 1, \dots, l+1$ be block upper triangular matrices with the following block entries:

   $$U_{ij} = A_{j-i},\ V_{ij} = B_{j-i},\ W_{ij} = \delta_{(i+1)j} I_n, \quad i = 1, \dots, l+1, j = i, \dots, l+1.$$

   Then the system (1.2) is solvable with $T_0 \in \mathbf{GL}(n, \mathbb{C})$ if and only for $l = \kappa(A, A)$ the pairs (U, W) and (V, W) are simultaneously similar.

2. $(A_0, ..., A_\ell) \in (\mathbb{C}^{n\times n})^{\ell+1}$ is simultaneously similar to a diagonal tuple $(B_0, ..., B_\ell) \in (\mathbb{C}^{n\times n})^{\ell+1}$, i.e. each B_i is a diagonal matrix, if and only if $A_0, ..., A_\ell$ are $\ell + 1$ commuting diagonable matrices: $A_i A_j = A_j A_i$ for $i, j = 0, ..., \ell$.

3. For $A_0, ..., A_\ell \in (\mathbb{C}^{n\times n})^{\ell+1}$ TFAE:

   - $(A_0, ..., A_\ell)$ is simultaneously similar to an upper triangular tuple $(B_0, ..., B_\ell) \in (\mathbb{C}^{n\times n})^{\ell+1}$.
   - For any $0 \le i < j \le \ell$ and $F \in \mathcal{A}(A_0, ..., A_\ell))$ the matrix $(A_i A_j - A_j A_i)F$ is nilpotent.

4. Let $\mathcal{X}_0 = \mathcal{A}(A_0, ..., A_\ell) \subset \mathbb{F}^{n\times n}$ and define recursively

   $$\mathcal{X}_k = \sum_{0\le i<j\le \ell} (A_i A_j - A_j A_i)\mathcal{X}_{k-1} \subset \mathbb{F}^{n\times n}), \quad k = 1, ..., .$$

Then $(A_0, ..., A_\ell)$ is simultaneously similar to an upper triangular tuple if and only if the following two conditions hold:

- $A_i \mathcal{X}_k \subset \mathcal{X}_k, \quad i = 0, ..., \ell, \ k = 0, ...,.$
- There exists $q \geq 1$ such that $\mathcal{X}_q = \{0\}$ and $\mathcal{X}_k$ is a strict subspace of $\mathcal{X}_{k-1}$ for $k = 1, ..., q$.

5. If $A_0, ..., A_\ell \in \mathbb{C}^{n \times n}$ commute then $(A_0, ..., A_\ell)$ is simultaneously similar to an upper triangular tuple $(B_0, ..., B_\ell)$.

3 Property L

The property L was introduced in studied in [MT52, MT55]. Most of the results of this section can be found in [MF80, Fri81, Frixx]. In this section we consider only square pencils are $A(x) = A_0 + A_1 x \in \mathbb{C}[x]^{n \times n}, A(x_0, x_1) \in \mathbb{C}[x_0, x_1]^{n \times n}$, where $A_1 \neq 0$.

Definitions:

A pencil $A(x) \in \mathbb{C}[x]^{n \times n}$ has **property** L if all the eigenvalues of $A(x_0, x_1)$ are linear functions. That is $\lambda_i(x_0, x_1) = \alpha_i x_0 + \beta_i x_1$ is an eigenvalue of $A(x_0, x_1)$ of multiplicity n_i for $i = 1, ..., m$, where

$$n = \sum_{i=1}^{m} n_i, \quad (\alpha_i, \beta_i) \neq (\alpha_j, \beta_j), \text{ for } 1 \leq i < j \leq m.$$

A pencil $A(x) = A_0 + A_1 x$ is called hermitian if A_1, A_2 are hermitian.

Facts:

1. For a pencil $A(x) = A_0 + xA_1 \in \mathbb{C}[x]^{n \times n}$ TFAE:
   - $A(x)$ has property L.
   - The eigenvalues of $A(x)$ are polynomials of degree 1 at most.

- The characteristic polynomial of $A(x)$ splits to linear factors over $\mathbb{C}[x]$.

- There is an ordering of the eigenvalues of A_0 and A_1, $a_1, ..., a_n$ and $b_1, ..., b_n$, respectively, such that the eigenvalues of $A_0x_0 + A_1x_1$ are $a_1x_0 + b_1x_1, ..., a_nx_0 + b_nx_1$.

2. A pencil in $A(x)$ has property L if one of the following conditions hold:

   - $A(x)$ is similar over $\mathbb{C}(x)$ to an upper triangular matrix $U(x) \in \mathbb{C}(x)^{n\times n}$.

   - $A(x)$ is strictly similar to an upper triangular pencil $U(x) = U_0 + U_1x$.

   - $A(x)$ is similar over $\mathbb{C}[x]$ to a diagonal matrix $B(x) \in \mathbb{C}[x]^{n\times n}$.

   - $A(x)$ is strictly similar to diagonal pencil.

3. If a pencil $A(x_0, x_1)$ has property L, then any two distinct eigenvalues are not tangent at any point of $\mathbb{C} \cup \infty$.

4. Assume that $A(x)$ is pointwise diagonable on $\mathbb{C}$. Then $A(x)$ has property L. Furthermore $A(x)$ is similar over $\mathbb{C}[x]$ to a diagonal pencil $B(x) = B_0 + B_1x$. Suppose furthermore that A_1 is diagonable, i.e. $A(x_0, x_1)$ is pointwise diagonable on $\mathbb{C}^2$. Then $A(x)$ is strictly similar to a diagonal pencil $B(x)$, i.e. A_0 and A_1 are commuting diagonable matrices.

5. Let $A(x) = A_0 + A_1x \in \mathbb{C}[x]^{n\times n}$ such that A_1 and A_2 are diagonable and $A_0A_1 \neq A_1A_0$. Then exactly one of the following conditions hold:

   - $A(x)$ is not diagonable exactly at the points $\zeta_1, ..., \zeta_p$, where $1 \leq p \leq n(n-1)$.

- For $n \geq 3$ $A(x) \in \mathbb{C}[x]^{n\times n}$ is diagonable exactly at the points $\zeta_1 = 0, ..., \zeta_q$ for some $q \geq 1$. (*We do not know if this condition is satisfied for some pencil.*)

6. Let $A(x) = A_0 + A_1 x$ be a hermitian pencil satisfying $A_0A_1 \neq A_1A_0$. Then there exists $2q$ distinct complex points $\zeta_1, \overline{\zeta}_1..., \zeta_q, \overline{\zeta}_q \in \mathbb{C}\backslash\mathbb{R}$, $1 \leq q \leq \frac{n(n-1)}{2}$ such that $A(x)$ is not diagonable if and only if $x \in \{\zeta_1, \overline{\zeta}_1, ..., \zeta_q, \overline{\zeta}_q\}$.

4 Simultaneous similarity classification I

This section outlines the setting for the classification of conjugacy classes of $l+1$ tuples $(A_0, A_1, \ldots, A_l) \in (\mathbb{C}^{n\times n})^{l+1}$ under the simultaneous similarity. This classification depends on certain standard notions in algebraic geometry that are explained briefly in this section. Consult for example with [Sch77]. More specific details are given in [Fri83], [Fri85] and [Fri86]. A detailed solution to the classification of conjugacy classes of $l+1$ tuples is outlined in the next section.

Definitions:

$\mathcal{X} \subset \mathbb{C}^N$ is called an **affine algebraic variety**, called here a **variety**, if it is the zero set of a finite number of polynomial equations in $\mathbb{C}^N$.

It is known that an intersection of a finite or infinite number of varieties is a variety, which can be an empty set. A finite union of varieties in $\mathbb{C}^N$ is a variety.

$\mathcal{X}$ is called irreducible if $\mathcal{X}$ does not decompose in a nontrivial way to a union of two varieties.

Every variety $\mathcal{X}$ is a finite nontrivial union of irreducible varieties, which are called the **irreducible** components of $\mathcal{X}$.

Let $\mathcal{X} \subset \mathbb{C}^N$ be an irreducible variety Then $\mathcal{X}$ is path-wise connected. $x \in \mathcal{X}$ is called a **regular** (**smooth**) point of $\mathcal{X}$ if in the neighborhood of this point

$\mathcal{X}$ is a complex manifold of a fixed dimension d, which is called the **dimension** of $\mathcal{X}$ and is denoted by dim $\mathcal{X}$.

For a reducible variety $\mathcal{Y} \subset \mathbb{C}^N$, the **dimension** of $\mathcal{Y}$, denoted by dim $\mathcal{Y}$, is the maximum dimension of its irreducible components.

The set of singular (non-smooth) points of $\mathcal{X}$ is denoted by $\mathcal{X}_s$. $\mathcal{X}_s$ is a proper subvariety of $\mathcal{X}$ and dim $\mathcal{X}_s <$ dim $\mathcal{X}$.

dim $\mathbb{C}^N = N$ and $(\mathbb{C}^N)_s = \emptyset$. $\emptyset$ is an irreducible variety of dimension -1 and for any $z \in \mathbb{C}^N$ the set $\{z\}$ is an irreducible variety of dimension 0.

A set $\mathcal{Z}$ is called a **quasi-irreducible** variety if there exists an nonempty irreducible variety $\mathcal{X}$ and a strict subvariety $\mathcal{Y} \subset \mathcal{X}$ such that $\mathcal{Z} = \mathcal{X}\backslash\mathcal{Y}$. Then $\mathcal{Z}$ is path-wise connected and its closure, denoted by Cl $(\mathcal{Z})$, is equal to $\mathcal{X}$. We define the dimension of $\mathcal{Z}$, denoted by dim $\mathcal{Z}$ to be equal to the dimension of $\mathcal{X}$. Thus Cl $(\mathcal{Z})\backslash\mathcal{Z}$ is a variety of dimension strictly less than the dimension of $\mathcal{Z}$.

The set of all regular points of irreducible variety $\mathcal{X}$, denoted by $\mathcal{X}_r := \mathcal{X}\backslash\mathcal{X}_s$, is a quasi-irreducible variety. Moreover $\mathcal{X}_r$ is a path-wise connected complex manifold of complex dimension dim $\mathcal{X}$.

A quasi-variety $\mathcal{Z}$ is called **regular** if $\mathcal{Z} \subset \mathcal{X}_r$.

A stratification of $\mathbb{C}^N$ is a decomposition of $\mathbb{C}^N$ to a finite disjoint union of $\mathcal{X}_1, \ldots, \mathcal{X}_p$ of regular quasi-irreducible varieties such that Cl $(\mathcal{X}_i)\backslash\mathcal{X}_i = \cup_{j\in\mathcal{A}_i}\mathcal{X}_j$ for some $\mathcal{A}_i \subset \{1, ..., p\}$ for $i = 1, \ldots, p$. (Cl $(\mathcal{X}_i) = \mathcal{X}_i \iff \mathcal{A}_i = \emptyset$.)

Denote by $\mathbb{C}[\mathbb{C}^N]$ the ring of polynomial in N variables with coefficients in $\mathbb{C}$.

Denote by $\mathcal{W}_{n,l+1,r+1}$ the finite dimensional vector space of multi-linear polynomials in $(l+1)n^2$ variables of degree $r+1$ at most. That is, the degree of each variable in any polynomial is at most 1. Then $N(n, l, r) :=$ dim $\mathcal{W}_{n,l+1,r+1} = \sum_{i=0}^{r+1}\binom{(l+1)n^2}{i}$. $\mathcal{W}_{n,l+1,r+1}$ has a standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_{N(n,l,r)}$ in $\mathcal{W}_{n,l+1,r+1}$ consisting of monomials in $(l+1)n^2$ variables of degree $r+1$ at most, arranged in a lexicographical order.

Let $\mathcal{X} \subset \mathbb{C}^N$ be a quasi-irreducible variety. Denote by $\mathbb{C}[\mathcal{X}]$ the restriction of all polynomials $f(x) \in \mathbb{C}[\mathbb{C}^N]$ to $\mathcal{X}$. We identify $f, g \in \mathbb{C}[\mathbb{C}^N]$ if $f - g$ vanishes on $\mathcal{X}$. It is known that $\mathbb{C}[\mathcal{X}]$ is an integral domain. Let $\mathbb{C}(\mathcal{X})$ denote the quotient field of $\mathbb{C}[\mathcal{X}]$. $\mathbb{C}[\mathcal{X}], \mathbb{C}(\mathcal{X})$ can be identified with $\mathbb{C}[\mathrm{Cl}\ (\mathcal{X})], \mathbb{C}(\mathrm{Cl}\ (\mathcal{X}))$ respectively.

A rational function $h \in \mathbb{C}(\mathcal{X})$ is called **regular** if h is defined everywhere in $\mathcal{X}$. A regular rational function on $\mathcal{X}$ is an analytic function.

The group $\mathbf{GL}(n, \mathbb{C})$ acts by conjugation on $(\mathbb{C}^{n\times n})^{l+1}$:

$(A_0, \ldots, A_l) \mapsto T(A_0, \ldots, A_l)T^{-1} := (TA_0T^{-1}, \ldots TA_lT^{-1})$ for any $(A_0, \ldots, A_l) \in (\mathbb{C}^{n\times n})^{l+1}$ and $T \in \mathbf{GL}(n, \mathbb{C})$. Denote by $\mathbf{A}$ the $l+1$ tuple $(A_0, \ldots, A_l) \in (\mathbb{C}^{n\times n})^{l+1}$ respectively. Let orb $(\mathbf{A}) := \{T\mathbf{A}T^{-1} : T \in \mathbf{GL}(n, \mathbb{C})\}$ be the **orbit** of $\mathbf{A}$ (under the action of $\mathbf{GL}(n, \mathbb{C})$). It is known that orb $(\mathbf{A})$ is a quasi-irreducible variety in $(\mathbb{C}^{n\times n})^{l+1}$.

Let $\mathcal{X} \subset (\mathbb{C}^{n\times n})^{l+1}$ be an irreducible quasi-variety. $\mathcal{X}$ is called **invariant** (under the action of $\mathbf{GL}(n, \mathbb{C})$) if $T\mathcal{X}T^{-1} = \mathcal{X}$ for all $T \in \mathbf{GL}(n, \mathbb{C})$. Equivalently, $\mathcal{X}$ is invariant if $\mathbf{A} \in \mathcal{X} \iff$ orb $(\mathbf{A}) \subset \mathcal{X}$.

Assume that $\mathcal{X}$ is an invariant irreducible quasi-variety. A rational function $h \in \mathbb{C}(\mathcal{X})$ is called **invariant** if h is the same value on any two points of a given orbit in $\mathcal{X}$, where h is defined. Denote by $\mathbb{C}[\mathcal{X}]^{\mathrm{inv}} \subset \mathbb{C}[\mathcal{X}]$ and $\mathbb{C}(\mathcal{X})^{\mathrm{inv}} \subset \mathbb{C}(\mathcal{X})$ the sub-domain of invariant polynomials and subfield of invariant functions respectively. Clearly, the quotient field of $\mathbb{C}[\mathcal{X}]^{\mathrm{inv}}$ is a subfield of $\mathbb{C}(\mathcal{X})^{\mathrm{inv}}$, and in some interesting case the quotient field of $\mathbb{C}[\mathcal{X}]^{\mathrm{inv}}$ is a strict subfield of $\mathbb{C}(\mathcal{X})^{\mathrm{inv}}$.

Facts:

1. Assume that $\mathcal{X} \subset (\mathbb{C}^{n\times n})^{l+1}$ be an irreducible quasi-variety. Then $\mathbb{C}[\mathcal{X}]^{\mathrm{inv}}$ and $\mathbb{C}(\mathcal{X})^{\mathrm{inv}}$ are finitely generated. That is there exists $f_1, \ldots, f_i \in \mathbb{C}[\mathcal{X}]^{\mathrm{inv}}$ and $g_1, \ldots, g_j \in \mathbb{C}(\mathcal{X})^{\mathrm{inv}}$ such that any polynomial in $\mathbb{C}[\mathcal{X}]^{\mathrm{inv}}$ is a polynomial in $f_1, \ldots, f_i$, and any rational function in $\mathbb{C}(\mathcal{X})^{\mathrm{inv}}$ is a ra-

tional function in $g_1, \ldots, g_j$.

2. (**Classification Theorem.**) Let $n \geq 2$ and $l \geq 0$ be fixed integer. Then there exists a stratification $\cup_{i=1}^{p} \mathcal{X}_i$ of $(\mathbb{C}^{n\times n})^{l+1}$ with the following properties. For each $\mathcal{X}_i$ there exist m_i regular rational functions $g_{1,i}, \ldots, g_{m_i,i} \in \mathbb{C}(\mathcal{X}_i)^{\mathrm{inv}}$ such that the values of $g_{j,i}$ for $j = 1, \ldots, m_i$ on any orbit in $\mathcal{X}_i$ determines this orbit uniquely.

   The rational functions $g_{1,i}, \ldots, g_{m_i,i}$ are the generators of $\mathbb{C}(\mathcal{X}_i)^{\mathrm{inv}}$ for $i = 1, \ldots, p$.

Examples:

1. Let $\mathcal{S}$ be an irreducible variety of scalar matrices $\mathcal{S} := \{A \in \mathbb{C}^{2\times 2} : A = \frac{\text{trace } A}{2} I_2\}$ and $\mathcal{X} := \mathbb{C}^{2\times 2} \backslash \mathcal{S}$ be a quasi-variety. Then $\dim \mathcal{X} = 4$, $\dim \mathcal{S} = 1$ and $\mathbb{C}^{2\times 2} = \mathcal{X} \cup \mathcal{S}$ is a stratification of $\mathbb{C}^{2\times 2}$.

2. Let $\mathcal{U} \subset (\mathbb{C}^{2\times 2})^2$ be the set of all pairs $(A, B) \in (\mathbb{C}^{2\times 2})^2$ which are simultaneously similar to a pair of upper triangular matrices. Then $\mathcal{U}$ is a variety given by the zero of the following polynomial:

   $$\mathcal{U} := \{(A,B) \in (\mathbb{C}^{2\times 2})^2 : \ (2\,\text{trace } A^2 - (\text{trace } A)^2)(2\,\text{trace } B^2 - (\text{trace } B)^2) - (2\,\text{trace } AB - \text{trace } A\,\text{trace } B)^2 = 0\}.$$

   Let $\mathcal{C} \subset \mathcal{U}$ be the variety of commuting matrices:

   $$\mathcal{C} := \{(A,B) \in (\mathbb{C}^{2\times 2})^2 : \ AB - BA = 0\}.$$

   Let $\mathcal{V}$ be the variety given by the zeros of the following three polynomials

   $$\mathcal{V} := \{(A,B) \in (\mathbb{C}^{2\times 2})^2 : \quad 2\,\text{trace } A^2 - (\text{trace } A)^2 = 2\,\text{trace } B^2 - (\text{trace } B)^2 = 2\,\text{trace } AB - \text{trace } A\,\text{trace } B = 0\}.$$

   Then $\mathcal{V}$ is the variety of all pairs (A, B) which are simultaneously similar a pair of the form $\left(\begin{bmatrix} \lambda & \alpha \\ 0 & \lambda \end{bmatrix}, \begin{bmatrix} \mu & \beta \\ 0 & \mu \end{bmatrix}\right)$. Hence $\mathcal{V} \subset \mathcal{C}$. Let $\mathcal{W} :=$

$\{A \in (\mathbb{C}^{2\times 2}) :\ 2\,\mathrm{trace}\ A^2 - (\mathrm{trace}\ A)^2 = 0\}$ and $\mathcal{S} \subset \mathcal{W}$ be defined as in the previous example. Define the following quasi-varieties in $(\mathbb{C}^{2\times 2})^2$:

$$\mathcal{X}_1 := (\mathbb{C}^{2\times 2})^2 \backslash \mathcal{U},\ \mathcal{X}_2 := \mathcal{U}\backslash\mathcal{C},\ \mathcal{X}_3 = \mathcal{C}\backslash\mathcal{V},\ \mathcal{X}_4 := \mathcal{V}\backslash(\mathcal{S}\times\mathcal{W}\cup\mathcal{W}\times\mathcal{S}),$$
$$\mathcal{X}_5 := \mathcal{S}\times(\mathcal{W}\backslash S),\ \mathcal{X}_6 := (\mathcal{W}\backslash\mathcal{S})\times\mathcal{S},\ \mathcal{X}_7 = \mathcal{S}\times\mathcal{S}.$$

Then

$$\dim\ \mathcal{X}_1 = 8,\ \dim\ \mathcal{X}_2 = 7,\ \dim\ \mathcal{X}_3 = 6,\ \dim\ \mathcal{X}_4 = 5,$$
$$\dim\ \mathcal{X}_5 = \dim\ \mathcal{X}_6 = 4,\ \dim\ \mathcal{X}_7 = 2,$$

and $\cup_{i=1}^{7}\mathcal{X}_i$ is a stratification of $(\mathbb{C}^{2\times 2})^2$.

3. In the classical case of similarity classes in $\mathbb{C}^{n\times n}$, i.e. $l = 0$, it is possible to choose a fixed set of polynomial invariant functions as $g_j(A) = \mathrm{trace}\ (A^j)$ for $j = 1,\dots,n$. However we still have to stratify $\mathbb{C}^{n\times n}$ to $\cup_{i=1}^{p}\mathcal{X}_i$, where each $A \in \mathcal{X}_i$ has some specific Jordan structures.

4. Consider the stratification $\mathbb{C}^{2\times 2} = \mathcal{X}\cup\mathcal{S}$ as in the the first part of the Example *3*. Clearly $\mathcal{X}$ and $\mathcal{S}$ are invariant under the action of $\mathbf{GL}(2,\mathbb{C})$. The invariant functions $\mathrm{trace}\ A, \mathrm{trace}\ A^2$ determine uniquely $\mathrm{orb}\,(A)$ on $\mathcal{X}$. The Jordan canonical for of any A in $\mathcal{X}$ is either consists of two distinct Jordan blocks of order 1 or one Jordan block of order 2. The invariant function $\mathrm{trace}\ A$ determines $\mathrm{orb}\,(A)$ for any $A \in \mathcal{S}$. It is possible to refine the stratification of $\mathbb{C}^{2\times 2}$ to three invariant components $\mathbb{C}^{2\times 2}\backslash\mathcal{W}, \mathcal{W}\backslash\mathcal{S}, \mathcal{S}$, where $\mathcal{W}$ is defined in the second part of Example *3*. Each component contains only matrices with one kind of Jordan blocks. On the first component $\mathrm{trace}\ A, \mathrm{trace}\ A^2$ determine the orbit and on the second and third component $\mathrm{trace}\ A$ determines the orbit.

5. To see the fundamental difference between similarity ($l = 0$) and simultaneous similarity $l \geq 1$ it is suffices to consider the second part of Example

2. Observe first that the stratification of $(\mathbb{C}^{2\times 2})^2 = \cup_{i=1}^7 \mathcal{X}_i$ is invariant under the action of $\mathbf{GL}(2,\mathbb{C})$. On $\mathcal{X}_1$ the five invariant polynomials trace A, trace A^2, trace B, trace B^2, trace AB, which are algebraically independent, determine uniquely any orbit in $\mathcal{X}_1$.

Let $(A = (a_{ij})_{i,j=1}^2, B = (b_{ij})_{i,j=1}^2) \in \mathcal{X}_2$. Then A and B has a unique one dimensional common eigenspace corresponding to the eigenvalues λ_1, μ_1 of A, B respectively. Assume that $a_{12}b_{21} - a_{21}b_{12} \neq 0$. Define

$$\lambda_1 = \alpha(A,B) := \frac{(b_{11}-b_{22})a_{12}a_{21} + a_{22}a_{12}b_{21} - a_{11}a_{21}b_{12}}{a_{12}b_{21} - a_{21}b_{12}},$$
$$\mu_1 = \alpha(B,A).$$

Then trace A, trace B, $\alpha(A,B)$, $\alpha(B,A)$ are regular, algebraically independent, rational invariant functions on $\mathcal{X}_2$, whose values determine orb (A,B). Cl (orb (A,B)) contains an orbit generated by two diagonal matrices diag (λ_1,λ_2) and diag (μ_1,μ_2). Hence $\mathbb{C}[\mathcal{X}_2]^{\mathrm{inv}}$ is generated by the five invariant polynomials trace A, trace A^2, trace B, trace B^2, trace AB, which are algebraically dependent. Their values coincide exactly on two distinct orbits in $\mathcal{X}_2$. On $\mathcal{X}_3$ the above invariant polynomials separate the orbits.

Any $(A = [a_{ij}]_{i,j=1}^2, B = [a_{ij}]_{i,j=1}^2) \in \mathcal{X}_4$ is simultaneously similar a unique pair of the form $(\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \begin{bmatrix} \mu & t \\ 0 & \mu \end{bmatrix})$. Then $t = \gamma(A,B) := \frac{b_{12}}{a_{12}}$. Thus trace A, trace $B, \gamma(A,B)$ are three algebraically independent regular rational invariant functions on $\mathcal{X}_4$, whose values determine a unique orbit in $\mathcal{X}_4$. Clearly $(\lambda I_2, \mu I_2) \in$ Cl (X_4). Then $\mathbb{C}[\mathcal{X}_4]^{\mathrm{inv}}$ is generated by trace A, trace B. The values of trace $A = 2\lambda$, trace $B = 2\mu$ correspond to a complex line of orbits in $\mathcal{X}_4$. Hence the classification problem of simultaneous similarity classes in $\mathcal{X}_4$ or $\mathcal{V}$ is a *wild* problem.

On $\mathcal{X}_5, \mathcal{X}_6, \mathcal{X}_7$ the algebraically independent functions trace A, trace B

determine the orbit in each of the stratum.

5 Simultaneous similarity classification II

In this section we give an invariant stratification of $(\mathbb{C}^{n\times n})^{l+1}$, for $l \geq 1$, under the action of $\mathbf{GL}(n,\mathbb{C})$ and describe a set invariant regular rational functions on each stratum, which separate the orbits up to a finite number. We assume the nontrivial case $n > 1$. It is conjectured that the continuous invariants of the given orbit determines uniquely the orbit on each stratum given in the Classification Theorem.

Most of the results in this section are given in [Fri83]. Classification of simultaneous similarity classes of matrices is a known **wild** problem [GP69]. For another approach to classification of simultaneous similarity classes of matrices using **Belitskii** reduction see [Ser00]. See other applications of these techniques to classifications of linear systems [Fri85] and to canonical forms [Fri86].

Definitions:

For $\mathbf{A} = (A_0, \ldots, A_l), \mathbf{B} = (B_0, \ldots, B_l) \in (\mathbb{C}^{n\times n})^{l+1}$ let $L(\mathbf{B}, \mathbf{A}) : \mathbb{C}^{n\times n} \to (\mathbb{C}^{n\times n})^{l+1}$ be the linear operator given by $U \mapsto (B_0U - UA_0, \ldots, B_lU - UA_l)$. Then $L(\mathbf{B}, \mathbf{A})$ is represented by the $(l + 1)n^2 \times n^2$ matrix $(I_n \otimes B_0^{\mathrm{T}} - A_0 \otimes I_n, \ldots, I_n \otimes B_l^{\mathrm{T}} - A_l \otimes I_n)^{\mathrm{T}}$, where $U \mapsto (I_n \otimes B_0^{\mathrm{T}} - A_0 \otimes I_n, \ldots, I_n \otimes B_l^{\mathrm{T}} - A_l \otimes I_n)^{\mathrm{T}}U$. Let $L(\mathbf{A}) := L(\mathbf{A}, \mathbf{A})$. The dimension of orb $(\mathbf{A})$, denoted by dim orb $(\mathbf{A})$, is equal to the rank of $L(\mathbf{A})$.

Denote by $\mathcal{S}_n := \{A \in \mathbb{C}^{n\times n} : \ A = \frac{\text{trace } A}{n} I_n\}$ the variety of scalar matrices. Since any $U \in \mathcal{S}_n$ commutes with any $B \in \mathbb{C}^{n\times n}$ it follows that $\ker L(\mathbf{A}) \supset \mathcal{S}_n$. Hence rank $L(\mathbf{A}) \leq n^2 - 1$.

Let

$$\mathcal{M}_{n,l+1,r} := \{\mathbf{A} \in (\mathbb{C}^{n\times n})^{l+1} : \text{ rank } L(\mathbf{A}) = r\}, \quad r = 0, 1, \ldots, n^2 - 1.$$

Facts:

1. $\mathcal{M}_{n,l+1,n^2-1}$ is a invariant quasi-irreducible variety of dimension $(l+1)n^2$, i.e. $\mathrm{Cl}\,(\mathcal{M}_{n,l+1,n^2-1}) = (\mathbb{C}^{n\times n})^{l+1}$. All sets other $\mathcal{M}_{n,l+1,r}, r = n^2-2,\ldots,0$ have the decomposition to invariant quasi-irreducible varieties, each of dimension strictly less than $(l+1)n^2$.

2. Let $r \in [0, n^2-1]$, $\mathbf{A} \in \mathcal{M}_{n,l+1,r}$ and $\mathbf{B} = T\mathbf{A}T^{-1}$. Then $L(\mathbf{B},\mathbf{A}) = \mathrm{diag}\,(I_n \otimes T, \ldots, I_n \otimes T)L(\mathbf{A})(I_n \otimes T^{-1})$, rank $L(\mathbf{B},\mathbf{A}) = r$ and

   $\det L(\mathbf{B},\mathbf{A})[\alpha,\beta] = 0$ for any $\alpha \in \mathrm{Q}_{r+1,(l+1)n^2}, \beta \in \mathrm{Q}_{r+1,n^2}$.

3. Let $\mathbf{X} = (X_0,\ldots,X_l) \in (\mathbb{C}^{n\times n})^{l+1}$, with the indeterminate entries $X_k = (x_{k,ij})_{i,j=1}^n$ for $k = 0,\ldots,l$. Each $\det L(\mathbf{X},\mathbf{A})[\alpha,\beta], \alpha \in \mathrm{Q}_{r+1,(l+1)n^2}, \beta \in \mathrm{Q}_{r+1,n^2}$ is a vector in $\mathcal{W}_{n,l+1,r+1}$, i.e. it is a multi-linear polynomial in $(l+1)n^2$ variables of degree $r+1$ at most. We identify $\det L(\mathbf{X},\mathbf{A})[\alpha,\beta], \alpha \in \mathrm{Q}_{r+1,(l+1)n^2}, \beta \in \mathrm{Q}_{r+1,n^2}$ with the **row** vector $\mathbf{a}(\mathbf{A},\alpha,\beta) \in \mathbb{C}^{N(n,l,r)}$ given by its coefficients in the basis $\mathbf{e}_1,\ldots,\mathbf{e}_{N(n,l,r)}$. The number of these vectors is $M(n,l,r) := \binom{(l+1)n^2}{r+1}\binom{n^2}{r+1}$. Let $R(\mathbf{A}) \in \mathbb{C}^{M(n,l,r)N(n,l,r)\times M(n,l,r)N(n,l,r)}$ be the matrix with the rows $\mathbf{a}(A,\alpha,\beta)$, where the pairs $(\alpha,\beta) \in \mathrm{Q}_{r+1,(l+1)n^2} \times \mathrm{Q}_{r+1,n^2}$ are listed in a lexicographical order.

4. All points on the orb $(\mathbf{A})$ satisfies the following polynomial equations in $\mathbb{C}[(\mathbb{C}^{n\times n})^{l+1}]$;

   $$\det L(\mathbf{X},\mathbf{A})[\alpha,\beta] = 0, \tag{5.1}$$
   $$\texttt{for all}\ \alpha \in \mathrm{Q}_{r+1,(l+1)n^2}, \beta \in \mathrm{Q}_{r+1,n^2}.$$

   Thus the matrix $R(\mathbf{A})$ determines the above variety.

5. If $\mathbf{B} = T\mathbf{A}T^{-1}$ then $R(\mathbf{A})$ is **row equivalent** to $R(\mathbf{B})$. To each orb $(\mathbf{A})$ we can associate a unique **reduced row echelon form** $F(\mathbf{A}) \in$

$\mathbb{C}^{M(n,l,r)N(n,l,r)\times M(n,l,r)N(n,l,r)}$ of $R(\mathbf{A})$. $\rho(\mathbf{A}) := \mathrm{rank}\ R(\mathbf{A})$ is the number of linearly independent polynomials given in (5.1). Let

$\mathcal{I}(\mathbf{A}) = \{(1,j_1),\ldots,(\rho(\mathbf{A}),j_{\rho(\mathbf{A})})\} \subset \{1,\ldots,\rho(\mathbf{A})\} \times \{1,\ldots,N(n,l,r)\}$ be the location of the pivots in $F(\mathbf{A}) = (f_{ij}(\mathbf{A}))_{i,j=1}^{M(n,l,r),N(n,l,r)}$. That is $1 \le j_1 < \ldots < j_{\rho(\mathbf{A})} \le N(n,l,r)$, $f_{ij_i}(\mathbf{A}) = 1$ for $i = 1,\ldots,\rho(\mathbf{A})$ and $f_{ij} = 0$ unless $j \ge i$ and $i \in [1,\rho(\mathbf{A})]$. The nontrivial entries $f_{ij}(\mathbf{A})$ for $j > i$ are rational functions in the entries of the $l+1$ tuple $\mathbf{A}$. Thus $F(\mathbf{B}) = F(\mathbf{A})$ for $\mathbf{B} \in \mathrm{orb}\ (\mathbf{A})$. The numbers $r(\mathbf{A}) := \mathrm{rank}\ L(\mathbf{A})$, $\rho(A)$ and the set $\mathcal{I}(\mathbf{A})$ are called the **discrete** invariants of orb $(\mathbf{A})$. The rational functions $f_{ij}(\mathbf{A}), i = 1,\ldots,\rho(\mathbf{A}), j = i+1,\ldots,N(n,l,r)$ are called the **continuous** invariants of orb $(\mathbf{A})$.

6. (**Classification Theorem for Simultaneous Similarity**

Let $l \ge 1, n \ge 2$ be integers. Fix an integer $r \in [0,n^2-1]$ and let $M(n,l,r), N(n,l,r)$ be the integers defined as above. Let

$\rho \in [0,\min(M(n,l,r),N(n,l,r))]$ and the set $\mathcal{I} = \{(1,j_1),\ldots,(\rho,j_\rho) \subset \{1,\ldots,\rho\} \times \{1,\ldots,N(n,l,r)\}, 1 \le j_1 < \ldots < j_\rho \le N(n,l,r)$ be given. Let $\mathcal{M}_{n,l+1.r}(\rho,\mathcal{I})$ be the set of all $\mathbf{A} \in (\mathbb{C}^{n\times n})^{l+1}$ such that $\mathrm{rank}\ L(\mathbf{A}) = r$, $\rho(\mathbf{A}) = \rho$ and $\mathcal{I}(\mathbf{A}) = \mathcal{I}$. Then $\mathcal{M}_{n,l+1.r}(\rho,\mathcal{I})$ is invariant quasi-variety under the action of $\mathbf{GL}(n,\mathbb{C})$. Suppose that $\mathcal{M}_{n,l+1.r}(\rho,\mathcal{I}) \ne \emptyset$. Recall that for each $\mathbf{A} \in \mathcal{M}_{n,l+1.r}(\rho,\mathcal{I})$ the continuous invariants of $\mathbf{A}$, which correspond to the entries $f_{ij}(\mathbf{A}),\ i = 1,\ldots,\rho,\ j = i+1,\ldots,N(n,l,r)$ of the reduced row echelon form of $R(\mathbf{A})$, are regular rational invariant functions on $\mathcal{M}_{n,l+1.r}(\rho,\mathcal{I})$. Then the values of the continuous invariants determines a finite number of orbits in $\mathcal{M}_{n,l+1.r}(\rho,\mathcal{I})$.

The quasi-variety $\mathcal{M}_{n,l+1.r}(\rho,\mathcal{I})$ decomposes uniquely as a finite union of invariant regular irreducible quasi-variety. The union of all these decompositions of $\mathcal{M}_{n,l+1.r}(\rho,\mathcal{I})$ for all possible values r,ρ and the sets $\mathcal{I}$

gives rise to an invariant stratification of $(\mathbb{C}^{n\times n})^{l+1}$.

References

[Fri78] S. Friedland, Extremal eigenvalue problems, *Bull. Brazilian Math. Soc.* 9 (1978), 13-40.

[Fri80] S. Friedland, Analytic similarities of matrices, *Lectures in Applied Math.*, Amer. Math. Soc. 18 (1980), 43-85 (edited by C.I. Byrnes and C.F. Martin).

[Fri81] S. Friedland, A generalization of the Motzkin-Taussky theorem, *Linear Algebra Appl.* 36 (1981), 103-109.

[Fri83] S. Friedland, Simultaneous similarity of matrices, *Advances in Mathematics*, 50 (1983), 189-265.

[Fri85] S. Friedland, Classification of linear systems, *Proc. of A.M.S. Conf. on Linear Algebra and Its Role in Systems Theory*, Contemp. Math. 47 (1985), 131-147.

[Fri86] S. Friedland, Canonical forms, *Frequency Domain and State Space Methods for Linear Systems*, 115-121, edited by C.I. Byrnes and A. Lindquist, North Holland, 1986.

[Frixx] S. Friedland, *Matrices*, a book in preparation.

[GB77] M.A. Gauger and C.I. Byrnes, Characteristic free, improved decidability criteria for the similarity problem, *Lin. Multilin. Alg.* 5 (1977), 153-158.

[GP69] I.M. Gelfand and V.A. Ponomarev, Remarks on classification of a pair of commuting linear transformation in a finite dimensional vector space, *Functional Anal. Appl.* 3 (1969), 325-326.

[GR65] R. Gunning and H. Rossi, *Analytic Functions of Several Complex Variables*, Prentice-Hall, New Jersey, 1965.

[Gur80] R.M. Guralnick, A note on the local-global prinicple for similarity of matrices, *Lin. ALg. Appl.* 30 (1980), 651-654.

[MM64] M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Prindle, Weber & Schmidt, Boston, 1964.

[MF80] N. Moiseyev and S. Friedland, The association of resonance states with incomplete spectrum of finite complex scaled Hamiltonian matrices, *Phys. Rev.* A 22 (1980), 619-624.

[MT52] T.S. Motzkin and O. Taussky, Pairs of matrices with property L, *Trans. Amer. Math. Soc.* 73 (1952), 108-114.

[MT55] T.S. Motzkin and O. Taussky, Pairs of matrices with property L. II, *Trans. Amer. Math. Soc.* 80 (1955), 387-401.

[Sch77] I.R. Schafarevich, *Basic Algebraic Geometry*, Springer-Verlag, 1977.

[Ser00] V.V. Sergeichuk, Canonical matrices for linear matrix problems, *Linear Algebra Appl.* 317 (2000), 53-102.

[Was63] W. Wasow, On holomorphically similar matrices, *J. Math. Anal. Appl.* 4 (1963), 202-206.

[Was77] W. Wasow, Arnold's canonical matrices and asymptotic simplifcation of ordinary differential equations, *Lin. Alg. Appl.* 18 (1977), 163-170.

[Was78] W. Wasow, *Topics in Theory of Linear Differential Equations Having Singularities with respect to a Parameter*, IRMA, Univ. L. Paster, Strassbourg, 1978.

[Wei67] K. Weierstrass, Zur theorie der bilinearen un quadratischen formen, *Monatsch. Akad. Wiss. Berlin*, 310-338, 1867.