

An informal introduction to perturbations of matrices determined up to similarity or congruence

Lena Klimenko

National Technical University of Ukraine “Kyiv Polytechnic Institute”, Prospect Peremogy 37, Kiev, Ukraine.

E-mail address: e.n.klimenko@gmail.com

Vladimir V. Sergeichuk*

Institute of Mathematics, Tereshchenkivska 3, Kiev, Ukraine.

E-mail address: sergeich@imath.kiev.ua

Abstract. The reductions of a square complex matrix A to its canonical forms under transformations of similarity, congruence, or $*$ congruence are unstable operations: these canonical forms and reduction transformations depend discontinuously on the entries of A . We survey results about their behavior under perturbations of A and about normal forms of all matrices $A + E$ in a neighborhood of A with respect to similarity, congruence, or $*$ congruence. These normal forms are called miniversal deformations of A ; they are not uniquely determined by $A + E$, but they are simple and depend continuously on the entries of E .

1. Introduction

The purpose of this survey is to give an informal introduction into the theory of perturbations of a square complex matrix A determined up to transformations of similarity $S^{-1}AS$, or congruence S^TAS , or $*$ congruence S^*AS , in which S is nonsingular and $S^* := \bar{S}^T$.

2010 *Mathematics Subject Classification.* 15A21, 15A63, 47A07, 47A55.

Key words: similarity, congruence, $*$ congruence, perturbations, miniversal deformations, closure graphs.

* Supported by the FAPESP grant 2012/18139-2. The work was done while this author was visiting the University of São Paulo, whose hospitality is gratefully acknowledged.

The reduction of a matrix to its Jordan form is an unstable operation: both the Jordan form and a reduction transformation depend discontinuously on the entries of the original matrix. For example, the Jordan matrix $J_2(\lambda) \oplus J_2(\lambda)$ (we denote by $J_n(\lambda)$ the $n \times n$ upper-triangular Jordan block with eigenvalue λ) is reduced by arbitrarily small perturbations to matrices

$$\left[\begin{array}{cc|cc} \lambda & 1 & & \\ & \lambda & \varepsilon & \\ \hline & & \lambda & 1 \\ & & & \lambda \end{array} \right] \quad \text{or} \quad \left[\begin{array}{cc|cc} \lambda & 1 & \varepsilon & \\ & \lambda & \lambda & \\ \hline & & \lambda & 1 \\ & & & \lambda \end{array} \right], \quad \varepsilon \neq 0, \quad (1)$$

whose Jordan canonical forms are $J_3(\lambda) \oplus J_1(\lambda)$ or $J_4(\lambda)$, respectively. Therefore, if the entries of a matrix are known only approximately, then it is unwise to reduce it to its Jordan form.

Furthermore, when investigating a family of matrices close to a given matrix, then although each individual matrix can be reduced to its Jordan form, it is unwise to do so since in such an operation the smoothness relative to the entries is lost.

Let J be a Jordan matrix.

- (a) Arnold [1] (see also [2, 3]) constructed a *miniversal deformation* of J ; i.e., a simple normal form to which all matrices $J + E$ close to J can be reduced by similarity transformations that smoothly depend on the entries of E .
- (b) Boer and Thijsse [6] and, independently, Markus and Parilis [22] found each Jordan matrix J' for which there exists an arbitrary small matrix E such that $J + E$ is similar to J' . For example, if $J = J_2(\lambda) \oplus J_2(\lambda)$, then J' is either J , or $J_3(\lambda) \oplus J_1(\lambda)$, or $J_4(\lambda)$ with the same λ (see (1)).

Using (b), it is easy to construct for small n the *closure graph* G_n for similarity classes of $n \times n$ complex matrices; i.e., the Hasse diagram of the partially ordered set of similarity classes of $n \times n$ matrices that are ordered as follows: $a \leq b$ if a is contained in the closure of b . Thus, the graph G_n shows how the similarity classes relate to each other in the affine space of $n \times n$ matrices.

In Section 2.1 we give a sketch of constructive proof of Arnold's theorem about miniversal deformations of Jordan matrices, and in Sections 2.2–2.4 we consider closure graphs for similarity classes and similarity bundles. In Sections 3 and 4 we survey analogous results about perturbations of matrices determined up to congruence or *congruence.

We do not survey the well-developed theory of perturbations of matrix pencils [9, 10, 11, 15, 18, 19]; i.e., of matrix pairs (A, B) up to equivalence transformations (RAS, RBS) with nonsingular R and S .

All matrices that we consider are complex matrices.

2. Perturbations of matrices determined up to similarity

2.1. Arnold’s miniversal deformations of matrices under similarity. In this section, we formulate Arnold’s theorem about miniversal deformations of matrices under similarity and give a sketch of its constructive proof. Since each square matrix is similar to a Jordan matrix, it suffices to study perturbations of Jordan matrices.

For each Jordan matrix

$$J = \bigoplus_{i=1}^t (J_{m_{i1}}(\lambda_i) \oplus \dots \oplus J_{m_{ir_i}}(\lambda_i)), \quad m_{i1} \geq m_{i2} \geq \dots \geq m_{ir_i} \quad (2)$$

with $\lambda_i \neq \lambda_j$ if $i \neq j$, we define the matrix of the same size

$$J + D := \bigoplus_{i=1}^t \begin{bmatrix} J_{m_{i1}}(\lambda_i) + 0^\downarrow & 0^\downarrow & \dots & 0^\downarrow \\ 0^\leftarrow & J_{m_{i2}}(\lambda_i) + 0^\downarrow & \dots & \vdots \\ \vdots & \dots & \dots & 0^\downarrow \\ 0^\leftarrow & \dots & 0^\leftarrow & J_{m_{ir_i}}(\lambda_i) + 0^\downarrow \end{bmatrix} \quad (3)$$

in which

$$0^\leftarrow := \begin{bmatrix} * & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ * & 0 & \dots & 0 \end{bmatrix} \quad \text{and} \quad 0^\downarrow := \begin{bmatrix} 0 & \dots & 0 \\ \vdots & \dots & \vdots \\ 0 & \dots & 0 \\ * & \dots & * \end{bmatrix}$$

are blocks whose entries are zeros and stars.

The following theorem of Arnold [1, Theorem 4.4] is also given in [2, Section 3.3] and [3, §30].

Theorem 2.1 ([1]). *Let J be the Jordan matrix (2). Then all matrices $J + X$ that are sufficiently close to J can be simultaneously reduced by some transformation*

$$J + X \mapsto S(X)^{-1}(J + X)S(X), \quad \begin{array}{l} S(X) \text{ is analytic} \\ \text{at } 0 \text{ and } S(0) = I, \end{array} \quad (4)$$

to the form $J + D$ defined in (3) whose stars are replaced by complex numbers that depend analytically on the entries of X . The number of stars is minimal that can be achieved by transformations of the form (4), it is equal to the codimension of the similarity class of J .

The matrix (3) with independent parameters instead of stars is called a *miniversal deformation* of J (see formal definitions in [1], [2], or [3]).

The codimension of the similarity class of J is defined as follows. For each $A \in \mathbb{C}^{n \times n}$ and a small matrix $X \in \mathbb{C}^{n \times n}$,

$$\begin{aligned} (I - X)^{-1}A(I - X) &= (I + X + X^2 + \dots)A(I - X) \\ &= A + (XA - AX) + X(XA - AX) + X^2(XA - AX) + \dots \\ &= A + \underbrace{XA - AX}_{\text{small}} + \underbrace{X(I - X)^{-1}(XA - AX)}_{\text{very small}} \end{aligned}$$

and so the similarity class of A in a small neighborhood of A can be obtained by a very small deformation of the affine matrix space $\{A + XA - AX \mid X \in \mathbb{C}^{n \times n}\}$. (By the Lipschitz property [24], if A and B are close to each other and $B = S^{-1}AS$ with a nonsingular S , then S can be taken near I_n .)

The vector space

$$T(A) := \{XA - AX \mid X \in \mathbb{C}^{n \times n}\}$$

is the *tangent space* to the similarity class of A at the point A . The numbers

$$\dim_{\mathbb{C}} T(A), \quad \text{codim}_{\mathbb{C}} T(A) := n^2 - \dim_{\mathbb{C}} T(A) \quad (5)$$

are called the *dimension* and *codimension* of the similarity class of A .

Remark 2.1. The matrix (3) is the direct sum of t matrices that are not block triangular. But each Jordan matrix J is permutation similar to some *Weyr matrix* $J^\#$ with the following remarkable property: all commuting with $J^\#$ matrices are upper block triangular. Producing with (3) the same transformations of permutation similarity, Klimenko and Sergeichuk [19] obtained an upper block triangular matrix $J^\# + D^\#$, which is a miniversal deformation of $J^\#$.

Now we show sketchily how all matrices near J can be reduced to the form (3) by near-identity elementary similarity transformations; which explains the structure of the matrix (3).

Lemma 2.1. *Two matrices are similar if and only if one can be transformed to the other by a sequence of the following transformations (which are called elementary similarity transformations; see [25, Section 1.40]):*

- (i) *Multiplying column i by a nonzero $a \in \mathbb{C}$; then dividing row i by a .*
- (ii) *Adding column i multiplied by $b \in \mathbb{C}$ to column j ; then subtracting row j multiplied by b from row i .*
- (iii) *Interchanging columns i and j ; then interchanging rows i and j .*

Proof. Let A and B be similar; that is, $S^{-1}AS = B$. Write S as a product of elementary matrices: $S = E_1 E_2 \dots E_t$. Then

$$A \mapsto E_1^{-1} A E_1 \mapsto E_2^{-1} E_1^{-1} A E_1 E_2 \mapsto \dots \mapsto E_t^{-1} \dots E_2^{-1} E_1^{-1} A E_1 E_2 \dots E_t = B$$

is a desired sequence of elementary similarity transformations. □

Sketch of the proof of Theorem 2.1. Two cases are possible.

Case 1: $t = 1$. Suppose first that $J = J_3(0) \oplus J_2(0)$. Let

$$J + E = [b_{ij}]_{i,j=1}^5 := \left[\begin{array}{ccc|cc} \varepsilon_{11} & 1 + \varepsilon_{12} & \varepsilon_{13} & \varepsilon_{14} & \varepsilon_{15} \\ \varepsilon_{21} & \varepsilon_{22} & 1 + \varepsilon_{23} & \varepsilon_{24} & \varepsilon_{25} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} & \varepsilon_{34} & \varepsilon_{35} \\ \hline \varepsilon_{41} & \varepsilon_{42} & \varepsilon_{43} & \varepsilon_{44} & 1 + \varepsilon_{45} \\ \varepsilon_{51} & \varepsilon_{52} & \varepsilon_{53} & \varepsilon_{54} & \varepsilon_{55} \end{array} \right] \quad (6)$$

be any matrix near J (i.e., all ε_{ij} are small). We need to reduce it to the form

$$\left[\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ * & * & * & * & * \\ \hline * & 0 & 0 & 0 & 1 \\ * & 0 & 0 & * & * \end{array} \right], \quad (7)$$

in which the $*$'s are small complex numbers, by those transformations from Lemma 2.1 that are close to the identity transformation.

Dividing column 2 of (6) by $1 + \varepsilon_{12}$ and multiplying row 2 by $1 + \varepsilon_{12}$ (transformation (i)), we make $b_{12} = 1$. Since ε_{12} is small, this transformation is near-identity and the obtained matrix is near J . Some b_{ij} and ε_{ij} have been changed, but we use the same notation for them.

Subtracting column 2 (with $\varepsilon_{12} = 0$) multiplied by ε_{11} from column 1, we make $b_{11} = 0$; the inverse transformation of rows (which must be done by the definition of transformation (ii)) slightly changes row 2. Analogously, we make $b_{13} = b_{14} = b_{15} = 0$ subtracting column 2; the inverse transformations of rows slightly change row 2.

We obtain

$$[b_{ij}]_{i,j=1}^5 = \left[\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ \varepsilon_{21} & \varepsilon_{22} & 1 + \varepsilon_{23} & \varepsilon_{24} & \varepsilon_{25} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} & \varepsilon_{34} & \varepsilon_{35} \\ \hline \varepsilon_{41} & \varepsilon_{42} & \varepsilon_{43} & \varepsilon_{44} & 1 + \varepsilon_{45} \\ \varepsilon_{51} & \varepsilon_{52} & \varepsilon_{53} & \varepsilon_{54} & \varepsilon_{55} \end{array} \right]$$

with row 1 as in (7). In the same manner, we make $b_{23} = 1$ dividing column 3 by $1 + \varepsilon_{23}$, and then $b_{21} = b_{22} = b_{24} = b_{25} = 0$ subtracting column 3 (transformations (i) and (ii)); the inverse transformations with rows slightly change row 3. In the obtained matrix, we make $b_{45} = 1$; then $b_{41} = b_{42} = b_{43} = b_{44} = 0$; the inverse transformations with rows slightly change row 5.

We have obtained a matrix of the form

$$\left[\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ * & * & * & * & * \\ \hline 0 & 0 & 0 & 0 & 1 \\ * & * & * & * & * \end{array} \right] \quad (*\text{'s are small numbers})$$

by using near-identity elementary similarity transformations with (6).

To reduce the number of stars, we subtract row 2 multiplied by b_{53} from row 5 making $b_{53} = 0$. The inverse transformation of columns adds column 5 multiplied by the old b_{53} to column 2. Then we make $b_{42} = b_{52} = 0$ using row 1; the inverse transformations of columns slightly change b_{31} , b_{41} , and b_{51} .

We have simultaneously reduced all matrices (6) near J to the form (7) by a similarity transformation that analytically depends on all ε_{ij} and that is identity if all $\varepsilon_{ij} = 0$.

In the same manner, all matrices $J(0) + E$ near a nilpotent Jordan matrix

$$J(0) := J_{m_1}(0) \oplus \cdots \oplus J_{m_r}(0), \quad m_1 \geq m_2 \geq \cdots \geq m_r$$

can be reduced first to matrices of the form

$$\left[\begin{array}{ccc} J_{m_1}(0) + 0^\downarrow & \cdots & 0^\downarrow \\ \vdots & \ddots & \vdots \\ 0^\downarrow & \cdots & J_{m_r}(0) + 0^\downarrow \end{array} \right]$$

and then to matrices of the form (3) with $t = 1$, $\lambda_1 = 0$, and m_1, \dots, m_r instead of m_{11}, \dots, m_{1r_1} .

This proves the theorem for each Jordan matrix $J(\lambda) = J(0) + \lambda I$ with a single eigenvalue λ since $S(E)^{-1}(J(\lambda) + E)S(E) = S(E)^{-1}(J(0) + E)S(E) + \lambda I$.

Case 2: $t \geq 2$. In this case, (2) has distinct eigenvalues. Write (2) in the form $J = J_1 \oplus \cdots \oplus J_t$, where each $J_i := J_{m_{i1}}(\lambda_i) \oplus \cdots \oplus J_{m_{ir_i}}(\lambda_i)$ is of size $n_i \times n_i$ and has the single eigenvalue λ_i . Let

$$J + E = \begin{bmatrix} J_1 + E_{11} & \cdots & E_{1t} \\ \vdots & \ddots & \vdots \\ E_{t1} & \cdots & J_t + E_{tt} \end{bmatrix} \quad (8)$$

be any matrix near J (i.e., all E_{ij} are small). We make $E_{ij} = 0$ for all $i \neq j$ by near-identity similarity transformations as follows.

Represent (8) in the form $J + E^{\not\sim} + E^{\not\approx}$ in which

$$J + E^{\not\sim} := \begin{bmatrix} J_1 & & & 0 \\ E_{21} & J_2 & & \\ \vdots & \ddots & \ddots & \\ E_{t1} & \dots & E_{t,t-1} & J_t \end{bmatrix}, \quad E^{\not\approx} := \begin{bmatrix} E_{11} & E_{12} & \dots & E_{1t} \\ & E_{22} & \ddots & \vdots \\ & & \ddots & E_{t-1,t} \\ 0 & & & E_{tt} \end{bmatrix}.$$

Let us reduce $J + E^{\not\sim}$. Add to its first vertical strip the second strip multiplied by any $n_2 \times n_1$ matrix M to the right. Make the inverse transformation of rows: subtract from the second horizontal strip the first strip multiplied by M to the left. This similarity transformation replaces E_{21} with $E_{21} + J_2M - MJ_1$. Since J_1 and J_2 have distinct eigenvalues, there exists M for which $E_{21} + J_2M - MJ_1 = 0$ (see [14, Chapter VIII, §3]). Moreover, M is small since E_{21} is small.

In the same manner, we successively make zero the other blocks of the first underdiagonal $E_{21}, E_{32}, \dots, E_{t,t-1}$ of $J + E^{\not\sim}$, then the blocks of its second underdiagonal $E_{31}, \dots, E_{t,t-2}$, and so on. Thus, there exists a near-identity matrix S_1 such that $S_1^{-1}(J + E^{\not\sim})S_1 = J_1 \oplus \dots \oplus J_t$.

We make the same similarity transformation with the whole matrix $J + E = J + E^{\not\sim} + E^{\not\approx}$ and obtain the matrix $J + E' := S_1^{-1}(J + E)S_1$. Its underdiagonal blocks E'_{ij} ($i > j$) coincide with the underdiagonal blocks of $S_1^{-1}E^{\not\approx}S_1$, which are very small since all E_{ij} are small and the transformation is near-identity.

We apply the same reduction to $J + E'$ and obtain a matrix $J + E'' = S_2^{-1}(J + E')S_2$ whose underdiagonal blocks E''_{ij} ($i > j$) are very very small, and so on.

The infinite product $S_1S_2\dots$ converges to a near-identity matrix S such that all underdiagonal blocks of $J + \tilde{E} := S^{-1}(J + E)S$ are zero.

By near-identity similarity transformations, we successively make zero the first overdiagonal $\tilde{E}_{12}, \tilde{E}_{23}, \dots, \tilde{E}_{t-1,t}$ of $J + \tilde{E}$, then its second overdiagonal $\tilde{E}_{13}, \dots, \tilde{E}_{t-2,t}$, and so on.

We have reduced (8) to the block diagonal form $(J_1 + F_1) \oplus \dots \oplus (J_t + F_t)$ in which all F_i are small. Reducing each summand $J_i + F_i$ as in Case 1, we obtain a matrix of the form (3). \square

Remark 2.2. In the above proof we have described sketchily how to construct the transformation (4). Algorithms for constructing this transformation are discussed in [20, 21].

2.2. Change of the Jordan canonical form by arbitrarily small perturbations. Let J be a Jordan matrix and let λ be its eigenvalue. Denote

by w_{λ_j} the number of Jordan blocks $J_m(\lambda)$ of size $m \geq j$ in J ; the sequence $(w_{\lambda_1}, w_{\lambda_2}, \dots)$ is called the *Weyr characteristic* of J (and of any matrix that is similar to J) for the eigenvalue λ .

The following theorem was proved by Boer and Thijsse [6] and, independently, by Markus and Parilis [22]; another proof was given by Elmroth, Johansson, and Kågström [10, Theorem 2.2].

Theorem 2.2 ([6, 22]). *Let J and J' be Jordan matrices of the same size. Then J can be transformed to a matrix that is similar to J' by an arbitrarily small perturbation if and only if J and J' have the same set of eigenvalues with the same multiplicities, and their Weyr characteristics satisfy*

$w_{\lambda_1} \geq w'_{\lambda_1}$, $w_{\lambda_1} + w_{\lambda_2} \geq w'_{\lambda_1} + w'_{\lambda_2}$, $w_{\lambda_1} + w_{\lambda_2} + w_{\lambda_3} \geq w'_{\lambda_1} + w'_{\lambda_2} + w'_{\lambda_3}$, \dots
for each eigenvalue λ .

Theorem 2.2 was extended to Kronecker's canonical forms of matrix pencils by Pokrzywa [23].

2.3. Closure graphs for similarity classes.

Definition 2.1. Let T be a topological space with an equivalence relation. The *closure graph* (or *closure diagram*) is the directed graph whose vertices bijectively correspond to the equivalence classes and for equivalence classes a and b there is a directed path from a vertex of a to a vertex of b if and only if $a \subset \bar{b}$, in which \bar{b} denotes the closure of b .

Thus, the closure graph is the Hasse diagram of the set of equivalence classes with the following partial order: $a \preceq b$ if and only if $a \subset \bar{b}$. The closure graph shows how the equivalence classes relate to each other in T .

In this section, $T = \mathbb{C}^{n \times n}$ and the equivalence relation is the similarity of matrices. Since each similarity class contains exactly one Jordan matrix determined up to permutations of Jordan blocks, we identify the vertices with the Jordan matrices determined up to permutations of Jordan blocks.

Theorem 2.2 admits to construct the closure graphs due to the following lemma.

Lemma 2.2. *The closure graph for similarity classes of $n \times n$ matrices contains a directed path from a Jordan matrix J to a Jordan matrix J' if and only if J can be transformed to a matrix that is similar to J' by an arbitrarily small perturbation.*

Proof. Denote by $[M]$ the similarity class of a square matrix M .

“ \Leftarrow ” Let J can be transformed to a matrix that is similar to J' by an arbitrarily small perturbation. Then there exists a sequence of matrices

$J + E_1, J + E_2, J + E_3, \dots$ in $[J']$ that converges to J . This means that $J \in \overline{[J']}$. Let $A \in [J]$; i.e., $A = S^{-1}JS$ for some S . Then the sequence of matrices $S^{-1}(J + E_i)S = A + S^{-1}E_iS$ ($i = 1, 2, \dots$) in $[J']$ converges to A , and so $A \in \overline{[J']}$. Therefore, $[J] \subset \overline{[J']}$ and there is a directed path from J to J' . \square

Corollary 2.1. *By Theorem 2.2, the arrows are only between Jordan matrices with the same sets of eigenvalues. Let J be a Jordan matrix.*

- *Let J' be a Jordan matrix of the same size. Each neighborhood of J contains a matrix whose Jordan canonical form is J' if and only if there is a directed path from J to J' (if $J = J'$ then there always exists the “lazy” path of length 0 from J to J').*
- *The closure of the similarity class of J is equal to the union of the similarity classes of all Jordan matrices J' such that there is a directed path from J' to J (if $J = J'$ then there always exists the “lazy” path).*

Example 2.1. Let us construct the closure graph for similarity classes of 4×4 matrices. Each Jordan matrix is a direct sum of Jordan blocks $J_m(\lambda)$. Replacing them by λ^m and deleting the symbols \oplus , we get the compact notation of Jordan matrices which was used by Arnold [1]. For example,

$$\lambda^2\lambda\mu \text{ is } J_2(\lambda) \oplus J_1(\lambda) \oplus J_1(\mu) \tag{9}$$

(we write λ, μ instead of λ^1, μ^1).

For all Jordan matrices of size 4×4 with eigenvalue 0, we have

Jordan matrix	its Weyr characteristic (w_1, w_2, w_3, w_4)	the sequence $(w_1, w_1 + w_2, w_1 + w_2 + w_3, w_1 + w_2 + w_3 + w_4)$
0000	(4,0,0,0)	(4,4,4,4)
0 ² 00	(3,1,0,0)	(3,4,4,4)
0 ² 0 ²	(2,2,0,0)	(2,4,4,4)
0 ³ 0	(2,1,1,0)	(2,3,4,4)
0 ⁴	(1,1,1,1)	(1,2,3,4)

Using this table, Theorem 2.2, and Lemma 2.2, it is easy to construct the following closure graph for similarity classes of nilpotent 4×4 matrices:

$$0000 \rightarrow 0^200 \rightarrow 0^20^2 \rightarrow 0^30 \rightarrow 0^4$$

In the same way, one can construct the closure graph for similarity classes of all 4×4 matrices, which is presented in Figure 1. The graph is infinite: λ, μ, ν, ξ are arbitrary distinct complex numbers. The similarity classes of 4×4 Jordan matrices J that are located at the same horizontal level in (10) have the same dimension (defined in (5)), which is indicated to the

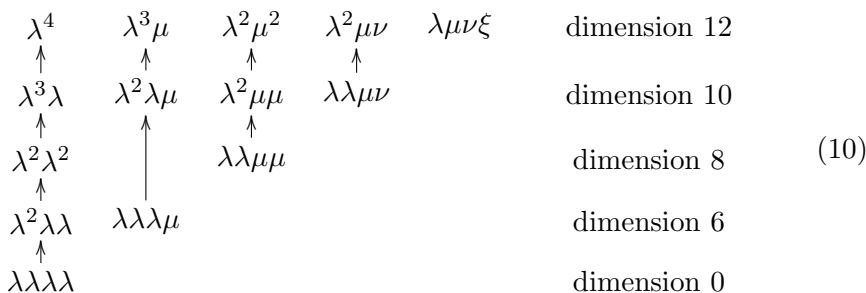


FIGURE 1. The closure graph for similarity classes of 4×4 matrices

right and is calculated as follows: it equals $16 - \text{codim}_{\mathbb{C}} T(J)$, in which $\text{codim}_{\mathbb{C}} T(J)$ is the number of stars in (3) (see (5) and Theorem 2.1). For example, if J is (9) with $\lambda \neq \mu$, then (3) is

$$\left[\begin{array}{cc|cc}
 \lambda & 1 & 0 & 0 \\
 * & \lambda + * & * & 0 \\
 * & 0 & \lambda + * & 0 \\
 0 & 0 & 0 & \mu + *
 \end{array} \right]$$

and so $\dim_{\mathbb{C}}(J) = 16 - \text{codim}_{\mathbb{C}} T(J) = 16 - 6 = 10$.

The following example shows that the structure of the closure graph for larger matrices is not so simple as in (10).

Example 2.2. The closure graph for similarity classes of 6×6 nilpotent matrices is presented in Figure 2. This graph was taken from [18, Figures 3 and 22], where P. Johansson describes the StratiGraph, which is a software tool for constructing the closure graphs for similarity classes of matrices, for strict equivalence classes of matrix pencils, and for bundles of matrices and pencils (see Section 2.4 about bundles and the web page

<http://www.cs.umu.se/english/research/groups/matrix-computations/stratigraph/>

about the StratiGraph).

2.4. Closure graphs for similarity bundles. Arnold [1, § 5.3] defines a *bundle of matrices under similarity* as a set of all matrices having the same *Jordan type*, which is defined as follows: matrices A and B have the same Jordan type if there is a bijection from the set of distinct eigenvalues of A to the set of distinct eigenvalues of B that transforms the Jordan canonical form of A to the Jordan canonical form of B . For example, the Jordan

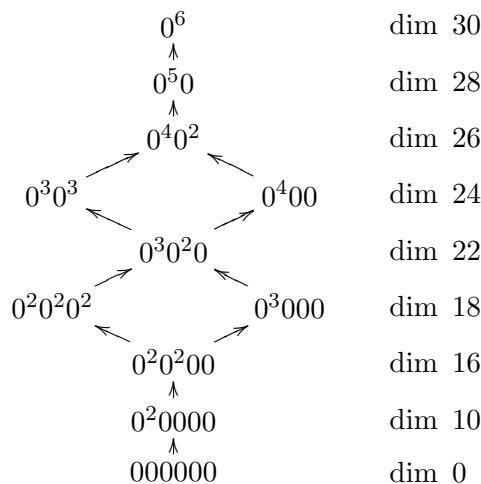


FIGURE 2. The closure graph for similarity classes of 6×6 nilpotent matrices

matrices

$$J_3(0) \oplus J_2(0) \oplus J_5(1), \quad J_3(2) \oplus J_2(2) \oplus J_5(-3)$$

belong to the same bundle. All matrices of a bundle have similar properties and not only with respect to perturbations; for example, its Jordan canonical matrices have the same set of commuting matrices.

Note that the closure graph for bundles of $n \times n$ matrices under similarity has a finite number of vertices; moreover, it is in some sense more informative than the closure graph for similarity classes. For example, one cannot see from the latter graph that each neighborhood of $J_n(\lambda)$ contains a matrix with n distinct eigenvalues (since there is no diagonal matrix whose similarity class has a nonzero intersection with each neighborhood of $J_n(\lambda)$). But the closure graph for bundles has an arrow from the bundle containing $J_n(\lambda)$ to the bundle of all matrices with n distinct eigenvalues.

Furthermore, not every convergent sequence of $n \times n$ matrices

$$B_1, B_2, \dots \rightarrow A, \tag{11}$$

in which all B_i are not similar to A , gives a directed path in the closure graph for similarity classes. But every sequence (11), in which all B_i do not belong to the bundle A that contains A , gives at least one directed path in the closure graph for similarity bundles. Indeed, the number of bundles of $n \times n$ matrices is finite, and so there is an infinite subsequence

$B_{n_1}, B_{n_2}, \dots \rightarrow A$ in which all B_{n_i} belong to the same bundle B . Hence $A \in \overline{B}$. One can prove that $A \subset \overline{B}$.

Example 2.3. The closure graph for similarity bundles of 4×4 matrices is presented in Figure 3 (it is given in another form in Johansson’s guide [18, Figure 24]).

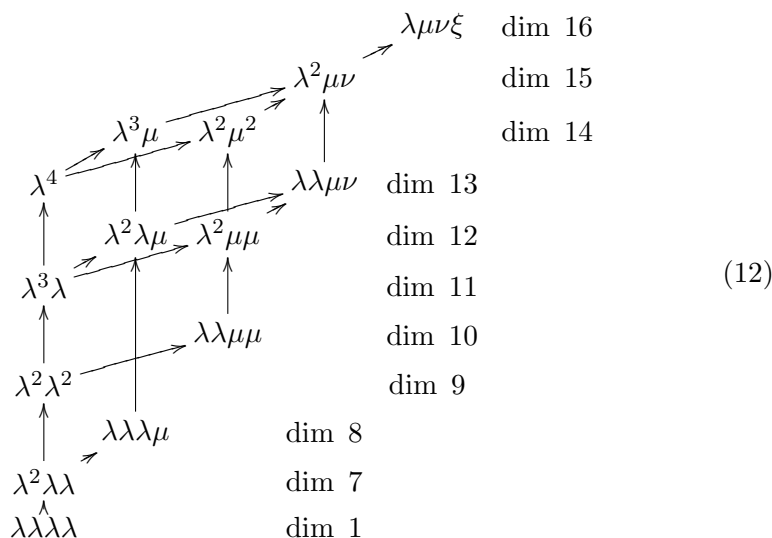


FIGURE 3. The closure graph for similarity bundles of 4×4 matrices

Let us compare (10) and (12). The graph (10) is infinite; it is the disjoint union of linear subgraphs that are obtained from

$$\lambda\lambda\lambda\lambda \rightarrow \lambda^2\lambda\lambda \rightarrow \dots \rightarrow \lambda^4, \quad \lambda\lambda\lambda\mu \rightarrow \lambda^2\lambda\mu \rightarrow \lambda^3\mu, \dots, \lambda\mu\nu\xi \quad (13)$$

by replacing their parameters by unequal complex numbers (the numbers of parameters in the vertices of the linear subgraphs (13) are equal to 1, 2, 2, 3, 4, respectively). Thus, although the sequences of Greek letters in the vertices of (10) and (12) are the same, each vertex of (10) represents an infinite set of similarity classes whose matrices have the same Jordan type (and so these similarity classes have the same dimension), whereas the corresponding vertex in (12) represents only one bundle, which is the union of these similarity classes; its dimension is equal to the dimension of any of its similarity classes plus the number of parameters. Notice that each arrow of (10) corresponds to an arrow of (12), but (12) has additional arrows.

3. Perturbations of matrices determined up to congruence

Dmytryshyn, Futorny, and Sergeichuk [7] constructed miniversal deformations of the following congruence canonical matrices given by Horn and Sergeichuk [16, 17]:

Every square complex matrix is congruent to a direct sum, determined uniquely up to permutation of summands, of matrices of the form

$$\begin{bmatrix} 0 & I_m \\ J_m(\lambda) & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & & & -1 & \ddots \\ & 1 & & 1 & \ddots \\ & -1 & -1 & & \\ 1 & 1 & & & 0 \end{bmatrix}, \quad J_k(0),$$

in which $\lambda \in \mathbb{C} \setminus \{0, (-1)^{m+1}\}$ and is determined up to replacement by λ^{-1} .

The miniversal deformations [7, Theorem 2.2] of congruence canonical matrices are rather cumbersome, so we give them only for 2×2 and 3×3 matrices.

Theorem 3.1 ([7, Example 2.1]). *Let A be any 2×2 or 3×3 matrix. Then all matrices $A + X$ that are sufficiently close to A can be simultaneously reduced by some transformation*

$$S(X)^T(A + X)S(X), \quad S(X) \text{ is holomorphic at } 0, \quad (14)$$

to one of the following forms, in which $\lambda \in \mathbb{C} \setminus \{-1, 1\}$ and each nonzero λ is determined up to replacement by λ^{-1} .

- *If A is 2×2 :*

$$\begin{bmatrix} 0 & \\ & 0 \end{bmatrix} + \begin{bmatrix} * & * \\ * & * \end{bmatrix}, \quad \begin{bmatrix} 1 & \\ & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ * & * \end{bmatrix}, \quad \begin{bmatrix} 1 & \\ & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ * & 0 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} * & 0 \\ * & * \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} * & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ \lambda & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ * & 0 \end{bmatrix}.$$

- *If A is 3×3 :*

$$\begin{bmatrix} 0 & & \\ & 0 & \\ & & 0 \end{bmatrix} + \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}, \quad \begin{bmatrix} 1 & & \\ & 0 & \\ & & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ * & * & * \\ * & * & * \end{bmatrix},$$

$$\begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ * & 0 & 0 \\ * & * & * \end{bmatrix}, \quad \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ * & 0 & 0 \\ * & * & 0 \end{bmatrix},$$

$$\begin{aligned}
& \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} + \begin{bmatrix} * & 0 & 0 \\ * & * & 0 \\ * & * & * \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & \\ \lambda & 0 & \\ & & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ * & 0 & 0 \\ * & * & * \end{bmatrix} \quad (\lambda \neq 0), \\
& \begin{bmatrix} 0 & 1 & \\ 0 & 0 & \\ & & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ * & 0 & * \\ * & 0 & * \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 & \\ 1 & 1 & \\ & & 0 \end{bmatrix} + \begin{bmatrix} * & 0 & 0 \\ 0 & 0 & 0 \\ * & * & * \end{bmatrix}, \\
& \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 1 \end{bmatrix} + \begin{bmatrix} * & 0 & 0 \\ * & * & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & \\ \lambda & 0 & \\ & & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ * & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\
& \begin{bmatrix} 0 & -1 & \\ 1 & 1 & \\ & & 1 \end{bmatrix} + \begin{bmatrix} * & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ * & 0 & * \end{bmatrix}, \\
& \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & -1 \\ 1 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ * & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.
\end{aligned}$$

Each of these matrices has the form $A_{\text{can}} + D$ in which A_{can} is a canonical matrix for congruence and the stars in D are complex numbers that tend to zero as X tends to zero. The number of stars is the smallest that can be attained by using transformations (14); it is equal to the codimension of the congruence class of A .

The codimension of the congruence class of a congruence canonical matrix $A \in \mathbb{C}^{n \times n}$ was calculated by Dmytryshyn, Futorny, and Sergeichuk [7] and independently by De Terán and Dopico [4]; it is defined as follows. For each small matrix $X \in \mathbb{C}^{n \times n}$,

$$(I + X)^T A (I + X) = A + \underbrace{X^T A + AX}_{\text{small}} + \underbrace{X^T A X}_{\text{very small}}$$

and so the congruence class of A in a small neighborhood of A can be obtained by a very small deformation of the affine matrix space $\{A + X^T A + AX \mid X \in \mathbb{C}^{n \times n}\}$. (By the local Lipschitz property [24], if A and B are close to each other and $B = S^T A S$ with a nonsingular S , then S can be taken near I_n .)

The vector space

$$T(A) := \{X^T A + AX \mid X \in \mathbb{C}^{n \times n}\}$$

is the tangent space to the congruence class of A at the point A . The numbers

$$\dim_{\mathbb{C}} T(A), \quad \text{codim}_{\mathbb{C}} T(A) := n^2 - \dim_{\mathbb{C}} T(A)$$

are called the *dimension* and *codimension* of the congruence class of A .

Congruence bundles are defined by Futorny, Klimenko, and Sergeichuk [12] via bundles of matrix pairs under equivalence. Recall, that pairs (A, B) and (A', B') of $m \times n$ matrices are *equivalent* if there are nonsingular R and S such that $RAS = A'$ and $RBS = B'$. By Kronecker's theorem about matrix pencils [14, Chapter XII, §3], each pair (A, B) of matrices of the same size is equivalent to

$$L \oplus P_1(\lambda_1) \oplus \cdots \oplus P_t(\lambda_t), \quad \lambda_i \neq \lambda_j \text{ if } i \neq j, \quad \lambda_1, \dots, \lambda_t \in \mathbb{C} \cup \infty, \quad (15)$$

in which L is a direct sum of pairs of the form (L_k, R_k) and (L_k^T, R_k^T) , $k = 1, 2, \dots$, defined by

$$L_k := \begin{bmatrix} 1 & 0 & 0 \\ & \ddots & \ddots \\ 0 & & 1 & 0 \end{bmatrix}, \quad R_k := \begin{bmatrix} 0 & 1 & 0 \\ & \ddots & \ddots \\ 0 & & 0 & 1 \end{bmatrix} \quad ((k-1)\text{-by-}k),$$

and each $P_i(\lambda_i)$ is a direct sum of pairs of the form

$$(I_k, J_k(\lambda_i)) \text{ if } \lambda_i \in \mathbb{C} \quad \text{or} \quad (J_k(0), I_k) \text{ if } \lambda_i = \infty.$$

The direct sums L and $P_i(\lambda_i)$ are determined by (A, B) uniquely, up to permutation of summands. The *equivalence bundle* of (15) consists of all matrix pairs that are equivalent to pairs of the form

$$L \oplus P_1(\mu_1) \oplus \cdots \oplus P_t(\mu_t), \quad \mu_i \neq \mu_j \text{ if } i \neq j, \quad \mu_1, \dots, \mu_t \in \mathbb{C} \cup \infty,$$

with the same L, P_1, \dots, P_t (see [9]).

The definition of bundles of matrices under congruence is not so evident. They could be defined via the congruence canonical form by analogy with bundles of matrices under similarity and bundles of matrix pairs, but, unlike the Jordan and Kronecker canonical forms, the *perturbation behavior of a congruence canonical matrix with parameters depends on the values of its parameters*, which is illustrated by the canonical matrices $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ \lambda & 0 \end{bmatrix}$ in the left graph in Figure 4.

Definition 3.1 ([12]). Two square matrices A and B are in the same *congruence bundle* if and only if the pairs (A, A^T) and (B, B^T) are in the same equivalence bundle.

Definition 3.1 is based on Roiter's statement (see [12, Lemma 4.1]): two $n \times n$ matrices A and B are congruent if and only if the pairs (A, A^T) and (B, B^T) are equivalent.

Example 3.1. The closure graphs for congruence classes and congruence bundles of 2×2 matrices are presented in Figure 4; they were constructed by Futorny, Klimenko, and Sergeichuk [12].

The left graph: in Figure 4 is the closure graph for congruence classes of 2×2 matrices. The congruence classes are given by their

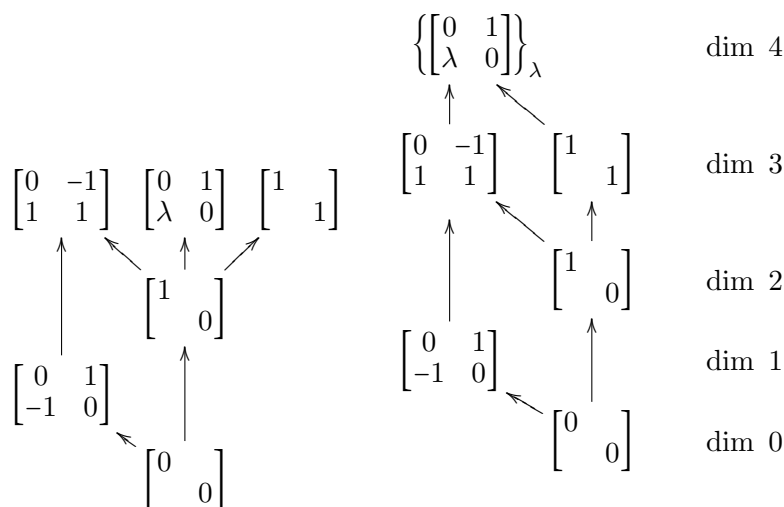


FIGURE 4. The closure graphs for congruence classes and congruence bundles of 2×2 matrices, in which $\lambda \in \mathbb{C} \setminus \{-1, 1\}$ and each nonzero λ is determined up to replacement by λ^{-1} .

2×2 canonical matrices for congruence. The graph is infinite: $\begin{bmatrix} 0 & 1 \\ \lambda & 0 \end{bmatrix}$ represents the infinite set of vertices indexed by $\lambda \in \mathbb{C} \setminus \{-1, 1\}$.

The right graph: is the closure graph for congruence bundles of 2×2 matrices. The vertex $\{\begin{bmatrix} 0 & 1 \\ \lambda & 0 \end{bmatrix}\}_\lambda$ represents the bundle that consists of all matrices whose congruence canonical forms are $\begin{bmatrix} 0 & 1 \\ \lambda & 0 \end{bmatrix}$ with $\lambda \neq \pm 1$. The other vertices are canonical matrices; their bundles coincide with their congruence classes. Note that $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ \lambda & 0 \end{bmatrix}$ ($\lambda \neq \pm 1$) properly belong to distinct bundles because these matrices have distinct properties with respect to perturbations, which is illustrated by the left graph. Other arguments in favor of Definition 3.1 of congruence bundles are given in [12, Section 6].

The congruence classes and bundles with vertices on the same horizontal level have the same dimension, which is indicated to the right.

Example 3.2. The closure graphs for congruence classes and congruence bundles of 3×3 matrices are presented in Figure 5. They were constructed by Futorny, Klimenko, and Sergeichuk [12].

The left graph: in Figure 5 is the closure graph for congruence classes of 3×3 matrices. The congruence classes are given by their 3×3 canonical matrices for congruence. The graph is infinite:

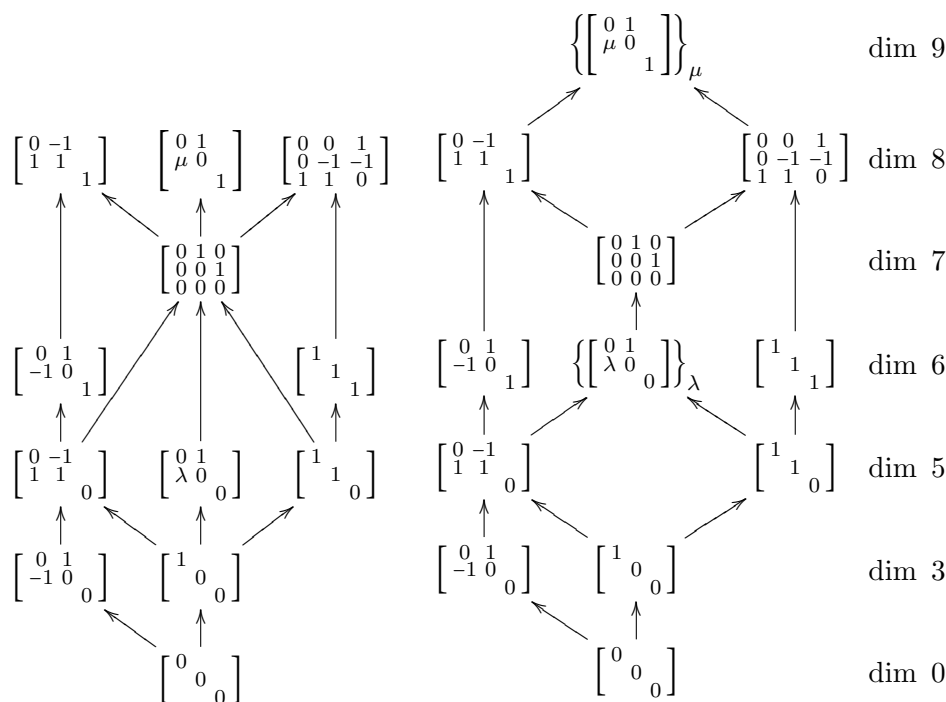


FIGURE 5. The closure graphs for congruence classes and congruence bundles of 3×3 matrices, in which $\lambda, \mu \neq \pm 1$, and nonzero λ and μ are determined up to replacements by λ^{-1} and μ^{-1} .

$\begin{bmatrix} 0 & 1 \\ \lambda & 0 \\ & & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ \mu & 0 \\ & & 1 \end{bmatrix}$ represent the infinite sets of vertices indexed by $\lambda, \mu \neq \pm 1$.

The right graph: is the closure graph for congruence bundles of 3×3 matrices. The vertices $\left\{ \begin{bmatrix} 0 & 1 \\ \lambda & 0 \\ & & 0 \end{bmatrix} \right\}_\lambda$ and $\left\{ \begin{bmatrix} 0 & 1 \\ \mu & 0 \\ & & 1 \end{bmatrix} \right\}_\mu$ represent the bundles that consist of all matrices whose congruence canonical forms are $\begin{bmatrix} 0 & 1 \\ \lambda & 0 \\ & & 0 \end{bmatrix}$ ($\lambda \neq \pm 1$) or $\begin{bmatrix} 0 & 1 \\ \mu & 0 \\ & & 1 \end{bmatrix}$ ($\mu \neq \pm 1$), respectively. The other vertices are canonical matrices; their bundles coincide with their congruence classes.

Remark 3.1. Let M be a 2×2 or 3×3 canonical matrix for congruence.

- Let N be another canonical matrix for congruence of the same size. Each neighborhood of M contains a matrix from the congruence

class (respectively, bundle) of N if and only if there is a directed path from M to N in the left (resp. right) graph in Figures 4 or 5. Note that there always exists the “lazy” path of length 0 from M to M if $M = N$.

- The closure of the congruence class (resp. bundle) of M is equal to the union of the congruence classes (resp. bundles) of all canonical matrices N such that there is a directed path from N to M .

4. Perturbations of matrices determined up *congruence

Dmytryshyn, Futorny, and Sergeichuk [8] constructed miniversal deformations of the following *congruence canonical matrices given by Horn and Sergeichuk [16, 17]:

*Every square complex matrix is *congruent to a direct sum, determined uniquely up to permutation of summands, of matrices of the form*

$$\begin{bmatrix} 0 & I_m \\ J_m(\lambda) & 0 \end{bmatrix}, \quad \mu \begin{bmatrix} 0 & & & 1 \\ & 1 & \ddots & i \\ & & \ddots & \\ 1 & i & & 0 \end{bmatrix}, \quad J_k(0), \quad (16)$$

in which $\lambda, \mu \in \mathbb{C}$, $|\lambda| > 1$, and $|\mu| = 1$. (The condition $|\lambda| > 1$ can be replaced by $0 < |\lambda| < 1$.)

The miniversal deformations [8, Theorem 2.2] of *congruence canonical matrices are rather cumbersome, so we give them only for 2×2 and 3×3 matrices.

Theorem 4.1. *Let A be any 2×2 or 3×3 matrix. Then all matrices $A + X$ that are sufficiently close to A can be simultaneously reduced by some transformation*

$$S(X)^*(A + X)S(X), \quad S(X) \text{ is nonsingular and continuous on a neighborhood of zero,}$$

to one of the following forms.

- *If A is 2×2 :*

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} * & * \\ * & * \end{bmatrix}, \quad \begin{bmatrix} \mu_1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 & 0 \\ * & * \end{bmatrix}, \quad \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 & 0 \\ \delta_{21} & \varepsilon_2 \end{bmatrix},$$

$$\begin{bmatrix} 0 & \mu_1 \\ \mu_1 & i\mu_1 \end{bmatrix} + \begin{bmatrix} * & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ \lambda & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ * & 0 \end{bmatrix}.$$

- *If A is 3×3 :*

$$\begin{array}{l}
 \begin{bmatrix} 0 & & \\ & 0 & \\ & & 0 \end{bmatrix} + \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}, \\
 \begin{bmatrix} \mu_1 & & \\ & \mu_2 & \\ & & 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 & 0 & 0 \\ \delta_{21} & \varepsilon_2 & 0 \\ * & * & * \end{bmatrix}, \\
 \begin{bmatrix} 0 & \mu_1 & \\ \mu_1 & i\mu_1 & \\ & & \mu_2 \end{bmatrix} + \begin{bmatrix} * & 0 & 0 \\ 0 & 0 & 0 \\ \delta_{21} & 0 & \varepsilon_2 \end{bmatrix}, \\
 \begin{bmatrix} 0 & 1 & \\ \lambda & 0 & \\ & & \mu_1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ * & 0 & 0 \\ 0 & 0 & \varepsilon_1 \end{bmatrix}, \\
 \begin{bmatrix} 0 & 1 & \\ 0 & 0 & \\ & & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ * & 0 & * \\ * & 0 & * \end{bmatrix}, \\
 \begin{bmatrix} 0 & 0 & \mu_1 \\ 0 & \mu_1 & i\mu_1 \\ \mu_1 & i\mu_1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \varepsilon_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\
 \begin{bmatrix} \mu_1 & & \\ & 0 & \\ & & 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 & 0 & 0 \\ * & * & * \\ * & * & * \end{bmatrix}, \\
 \begin{bmatrix} \mu_1 & & \\ & \mu_2 & \\ & & \mu_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 & 0 & 0 \\ \delta_{21} & \varepsilon_2 & 0 \\ \delta_{31} & \delta_{32} & \varepsilon_3 \end{bmatrix}, \\
 \begin{bmatrix} 0 & \mu_1 & \\ \mu_1 & i\mu_1 & \\ & & 0 \end{bmatrix} + \begin{bmatrix} * & 0 & 0 \\ 0 & 0 & 0 \\ * & * & * \end{bmatrix}, \\
 \begin{bmatrix} 0 & 1 & \\ \lambda & 0 & \\ & & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ * & 0 & 0 \\ * & * & * \end{bmatrix} \quad (\lambda \neq 0), \\
 \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ * & 0 & * \end{bmatrix},
 \end{array}$$

Each of these matrices has the form $A_{\text{can}} + D$, in which A_{can} is a canonical matrix for $*$ -congruence, the stars in D are complex numbers, $|\lambda| < 1$, $|\mu_1| = |\mu_2| = |\mu_3| = 1$, and

$$\begin{array}{ll}
 \varepsilon_l \in \mathbb{R} \text{ if } \mu_l \notin \mathbb{R} & \delta_{lr} = 0 \text{ if } \mu_l \neq \pm\mu_r \\
 \varepsilon_l \in i\mathbb{R} \text{ if } \mu_l \in \mathbb{R} & \delta_{lr} \in \mathbb{C} \text{ if } \mu_l = \pm\mu_r
 \end{array}$$

(Clearly, D tends to zero as X tends to zero.) For each $A_{\text{can}} + D$, twice the number of its stars plus the number of its entries $\varepsilon_l, \delta_{lr}$ is equal to the codimension over \mathbb{R} of the $*$ -congruence class of A_{can} .

The codimension of the $*$ -congruence class of a $*$ -congruence canonical matrix $A \in \mathbb{C}^{n \times n}$ was calculated by De Terán and Dopico [5] and independently by Dmytryshyn, Futorny, and Sergeichuk [8]; it is defined as follows. For each $A \in \mathbb{C}^{n \times n}$ and a small matrix $X \in \mathbb{C}^{n \times n}$,

$$(I + X)^* A (I + X) = A + \underbrace{X^* A + AX}_{\text{small}} + \underbrace{X^* A X}_{\text{very small}}$$

and so the $*$ -congruence class of A in a small neighborhood of A can be obtained by a very small deformation of the real affine matrix space $\{A + X^* A + AX \mid X \in \mathbb{C}^{n \times n}\}$. (By the local Lipschitz property [24], if A and B are close to each other and $B = S^* A S$ with a nonsingular S , then S can be taken near I_n). The real vector space

$$T(A) := \{X^* A + AX \mid X \in \mathbb{C}^{n \times n}\}$$

is the tangent space to the \ast -congruence class of A at the point A . The numbers

$$\dim_{\mathbb{R}} T(A), \quad \text{codim}_{\mathbb{R}} T(A) := 2n^2 - \dim_{\mathbb{R}} T(A)$$

are called the *dimension* and, respectively, *codimension over \mathbb{R}* of the \ast -congruence class of A .

Example 4.1. The closure graph for \ast -congruence classes of 2×2 matrices is presented in Figure 6; it was constructed by Futorny, Klimenko, and Sergeichuk [13]. Each \ast -congruence class is given by its canonical matrix,

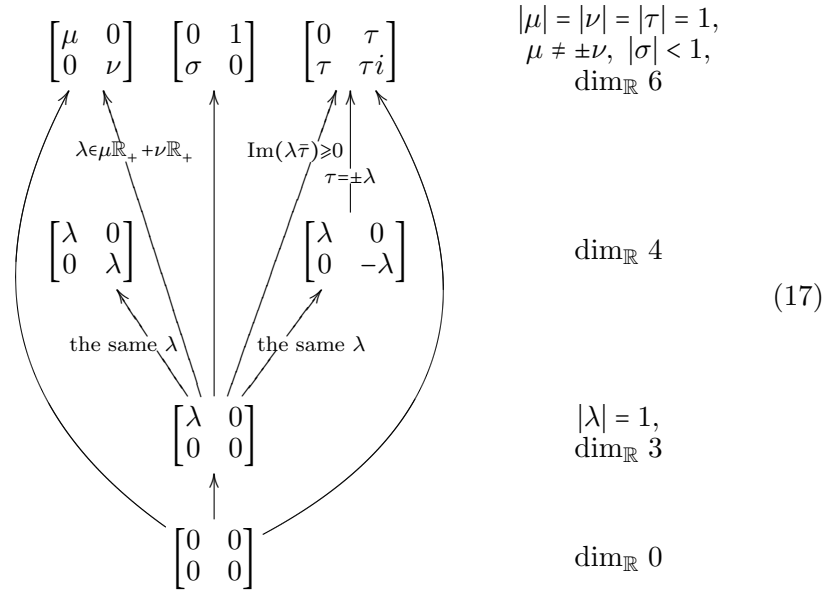


FIGURE 6. The closure graph for \ast -congruence classes of 2×2 matrices, in which \mathbb{R}_+ denotes the set of nonnegative real numbers, $\text{Im}(c)$ denotes the imaginary part of $c \in \mathbb{C}$, and $\lambda, \mu, \nu, \sigma, \tau \in \mathbb{C}$.

which is a direct sum of blocks of the form (16). The graph is infinite: each vertex except for $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ represents an infinite set of vertices indexed by the parameters of the corresponding canonical matrix. The \ast -congruence classes of canonical matrices that are located at the same horizontal level in (17) have the same dimension over \mathbb{R} , which is indicated to the right. The arrow $\begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} \mu & 0 \\ 0 & \nu \end{bmatrix}$ exists if and only if $\lambda = \mu a + \nu b$ for some nonnegative $a, b \in \mathbb{R}$. The arrow $\begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & \tau \\ \tau & i\tau \end{bmatrix}$ exists if and only if the imaginary part of $\lambda\bar{\tau}$ is nonnegative. The arrow $\begin{bmatrix} \lambda & 0 \\ 0 & -\lambda \end{bmatrix} \rightarrow \begin{bmatrix} 0 & \tau \\ \tau & i\tau \end{bmatrix}$ exists if and only if $\tau = \pm\lambda$.

The arrows $\begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} \lambda & 0 \\ 0 & \pm\lambda \end{bmatrix}$ exist if and only if the value of λ is the same in both matrices. The other arrows exist for all values of parameters of their matrices.

Remark 4.1. Let M be a 2×2 canonical matrix for *congruence.

- Let N be another 2×2 canonical matrix for *congruence. Each neighborhood of M contains a matrix that is *congruent to N if and only if there is a directed path from M to N in (17) (if $M = N$, then there always exists the “lazy” path of length 0 from M to N).
- The closure of the *congruence class of M is equal to the union of the *congruence classes of all canonical matrices N such that there is a directed path from N to M .

References

- [1] V.I. Arnold, On matrices depending on parameters, Russian Math. Surveys 26 (2) (1971) 29–43.
- [2] V.I. Arnold, Lectures on bifurcations in versal families, Russian Math. Surveys 27 (5) (1972) 54–123.
- [3] V.I. Arnold, Geometrical Methods in the Theory of Ordinary Differential Equations, Springer-Verlag, 1988.
- [4] F. De Terán, F.M. Dopico, The solution of the equation $XA + AX^T = 0$ and its application to the theory of orbits, Linear Algebra Appl. 434 (2011) 44–67.
- [5] F. De Terán, F.M. Dopico, The equation $XA + AX^* = 0$ and the dimension of *congruence orbits, Electr. J. Linear Algebra 22 (2011) 448–465.
- [6] H. den Boer, G.Ph.A. Thijssse, Semi-stability of sums of partial multiplicities under additive perturbation, Integral Equations Operator Theory 3 (1980) 23–42.
- [7] A.R. Dmytryshyn, V. Futorny, V.V. Sergeichuk, Miniversal deformations of matrices of bilinear forms, Linear Algebra Appl. 436 (2012) 2670–2700. (Preliminary version: Preprint RT-MAT 2007-04, Universidade de São Paulo, 2007, 34 p.)
- [8] A.R. Dmytryshyn, V. Futorny, V.V. Sergeichuk, Miniversal deformations of matrices under *congruence and reducing transformations, arXiv:1105.2160.
- [9] A. Edelman, E. Elmroth, B. Kågström, A geometric approach to perturbation theory of matrices and matrix pencils. Part I: Versal deformations, SIAM J. Matrix Anal. Appl. 18 (1997) 653–692.
- [10] A. Edelman, E. Elmroth, B. Kågström, A geometric approach to perturbation theory of matrices and matrix pencils. Part II: A stratification-enhanced staircase algorithm, SIAM J. Matrix Anal. Appl. 20 (1999) 667–699.
- [11] E. Elmroth, P. Johansson, B. Kågström, Computation and presentation of graph displaying closure hierarchies of Jordan and Kronecker structures, Numer. Linear Algebra Appl. 8 (2001) 381–399.
- [12] V. Futorny, L. Klimenko, V.V. Sergeichuk, Change of the congruence canonical form of 2-by-2 and 3-by-3 matrices under perturbations, arXiv:1004.3590.
- [13] V. Futorny, L. Klimenko, V.V. Sergeichuk, Change of the *congruence canonical form of 2-by-2 matrices under perturbations, arXiv:1304.5762.
- [14] F.R. Gantmacher, The Theory of Matrices, Vol. 1 and 2, Chelsea, New York, 1959.

- [15] M.I. Garcia-Planas, V.V. Sergeichuk, Simplest miniversal deformations of matrices, matrix pencils, and contragredient matrix pencils, *Linear Algebra Appl.* 302–303 (1999) 45–61.
- [16] R.A. Horn, V.V. Sergeichuk, Congruence of a square matrix and its transpose, *Linear Algebra Appl.* 389 (2004) 347–353.
- [17] R.A. Horn, V.V. Sergeichuk, Canonical forms for complex matrix congruence and *congruence, *Linear Algebra Appl.* 416 (2006) 1010–1032.
- [18] P. Johansson, *StratiGraph Users Guide*, Technical Report UMINF 03.21 (ISSN-0348-0542), Department of Computing Science, Umeå University, Sweden, 2003. Available at: <http://www8.cs.umu.se/pedher/research/papers/sg-usersguide.pdf>
- [19] L. Klimenko, V.V. Sergeichuk, Block triangular miniversal deformations of matrices and matrix pencils, in: V. Olshevsky, E. Tyrtyshnikov (Eds), *Matrix Methods: Theory, Algorithms and Applications*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2010, pp. 69–84.
- [20] A.A. Mailybaev, Transformation of families of matrices to normal forms and its application to stability theory, *SIAM J. Matrix Anal. Appl.* 21 (1999) 396–417.
- [21] A.A. Mailybaev, Transformation to versal deformations of matrices, *Linear Algebra Appl.* 337 (2001) 87–108.
- [22] A.S. Markus, E.È. Parilis, The change of the Jordan structure of a matrix under small perturbations, *Mat. Issled.* 54 (1980) 98–109. English translation: *Linear Algebra Appl.* 54 (1983) 139–152.
- [23] A. Pokrzywa, On perturbations and the equivalence orbit of a matrix pencil, *Linear Algebra Appl.* 82 (1986) 99–121.
- [24] L. Rodman, Remarks on Lipschitz properties of matrix groups actions, *Linear Algebra Appl.* 434 (2011) 1513–1524.
- [25] J.H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, 1965.