Universitat Politècnica de Catalunya
Departament de Matemàtiques

Stratification theory

of matrix pairs under equivalence and contragredient equivalence

by

Tetiana Klymchuk

Scientific advisor

Maria Isabel García-Planas

Barcelona – 2019

Pure mathematics is on the whole distinctly more useful than applied. For what is useful above all is technique, and mathematical technique is taught mainly through pure mathematics.

London 1941

G. H. Hardy

Acknowledgments

I would like to express my sincere gratitude to my advisor Prof. **Maria Isabel García-Planas** for the opportunity to spend 3 best years in my life in such a beautiful place like Barcelona; for the continuous support during my preparations and for the time that you spent preparing my visa—the reason which can be understood only by those who once tried to obtain a visa for non-European Union citizens; for your patience, motivation and support.

I thank every person that I had the opportunity to work side by side these years in Barcelona: **Marta Peña Carrera, Toni Susin** and **M. Dolors Magret Planas**. Thank you for your endless support in all the problems I had and for your caring during these years in Catalonia. With you, I have obtained my closest friends and I treat this like one of the most important achievements during my stay in Spain.

A special thanks to my family. Words cannot express how grateful I am for all of the sacrifices that you've made on my behalf. I certainly missed you a lot during these years. You were, are and will always be my support to everything. While I was studying, you provide maintaining in all areas in order to make my stay abroad safe and pleasant. Tremendous part of this degree belong to you and specially to my incredible mother **Svitlana Klymchuk**, my perfect grandparents **Fedir and Kateryna Chuprynda** and to my prototype and my personal hero, my grandma **Tamara Klymchuk**.

I really don't know the rights words to express my appreciation to my beloved husband **Dmytro Kovalenko** who was waiting for me from all

my scientific visits and conferences, who worked day and night when I was studying, who accepted my invitation to this life-changing adventure, which took us out from our comfort zone and who always understand me. Simply for everything, I thank you, I love you.

This thesis has been performed with the support of **Agència de Gestió d'Ajuts Universitaris i de Recerca** (FI-DGR 2016) from Generalitat de Catalunya, which co-financed by Fons Social Europeu.

CONTENT

INTRODUCTION

The theory of stratifications studies matrices and matrix pairs that are near each other. It has been largely studied by many authors: B. Kagstrom, A. Dmytryshyn, A. Edelman, E. Elmroth, S. Johansson [21, 23–25, 47], J. Ferrer [27], M.D. Magret [38]; M.I. García-Planas, V.V. Sergeichuk [39, 40], L. Klimenko [48], V. Futorny [20], and others. Fundamental concepts like controllability and observability are studied by methods of stratification theory.

If a matrix A is known only approximately, then there is no sense to reduce it to Jordan form, since it can be cardinally changed by a small perturbations. For example,

$$JCF\left[\begin{array}{cc|cc} \lambda & 1 & & \\ & \lambda & \varepsilon & \\ \hline & & \lambda & 1 \\ & & & \lambda \end{array}\right] = \begin{cases} \left[\begin{array}{cc|cc} \lambda & 1 & & \\ & \lambda & & \\ \hline & & \lambda & 1 \\ & & & \lambda \end{array}\right] & \text{if } \varepsilon = 0, \\ \begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix} & \text{if } \varepsilon \neq 0. \end{cases}$$

By this reason, it is important to know how the Jordan canonical form can change under small perturbations of A. Arnold [3] (see also [4, 5]) constructed miniversal deformations of matrices under similarity; that is, a simple normal form to which not only a given square matrix A but all matrices B close to it can be reduced by similarity transformations that smoothly depend on the entries of B.

Miniversal deformation of a pair (A, B) of complex matrices of the same size is a family of pairs $(A_{\text{def}}, B_{\text{def}})$ to which all matrix pairs (A', B') close to (A, B) can be reduced by transformations $(S^{-1}A'R, S^{-1}B'R)$ given by nonsingular matrices $S = S(A', B')$ and $R = R(A', B')$ that smoothly depend on the entries of A' and B'. This admits to study only miniversal deformations of a given matrix instead of considering all its perturbations.

We consider miniversal deformations of matrix pairs as a basic tool in the study of their stratification. The miniversal deformations of matrix pairs were obtained by Kagstrom, Edelman and Elmroth in the article [48], which was awarded by the SIAM Linear Algebra Prize 2000 for the most outstanding linear algebra publication during the 3-year period. We use the simpler miniversal deformations given by García-Planas and Sergeichuk [39].

Miniversal deformations have attracted the interest of the researchers in recent years due to the wide range of their applications. The study of behavior of a physical system that is described by matrix pencils require the understanding of how canonical structure may change, e.g., how eigenvalues coalesce or split apart due to perturbations. Sometimes it becomes an incredibly complicate task.

Canonical forms of matrices play important role both in theoretical and practical problems. Their analysis becomes very complicated, when we study multi-parameter families of matrices. We are faced with new phenomena like singularities and bifurcations leading to qualitative changes in the behavior of systems.

If we consider a couple of systems $y = Ax$ and $x = Bz$ in which $A \in M_{n\times m}(\mathbb{C})$ and $B \in M_{m\times p}(\mathbb{C})$, then we can write

$$\dot{y} = ABz.$$

If $p = n$ and $y = z$, the composite system corresponds to the classical homogeneous linear dynamical system $\dot{z} = ABz$:

$$z \longrightarrow \boxed{B} \xrightarrow{x} \boxed{A} \longrightarrow \dot{z} \qquad \text{implies} \qquad z \longrightarrow \boxed{AB} \longrightarrow \dot{z}$$

For a given matrix A, Boer and Thijsse [13] and, independently, Markus and Parilis [57] described the set of all Jordan canonical matrices J such that for each J there exists a matrix that is arbitrarily close to A and is similar to J. For example, if $A = J_2(\lambda) \oplus J_2(\lambda)$, then J is either $J_2(\lambda) \oplus J_2(\lambda)$, $J_3(\lambda) \oplus J_1(\lambda)$, or $J_4(\lambda)$ with the same λ.

Using their description, one can construct the *closure graph* G_n for similarity classes of $n \times n$ complex matrices; that is, the Hasse diagram of the set of similarity classes of $n \times n$ matrices with the following partial order: $a \preccurlyeq b$ if a is contained in the closure of b. The graph G_n shows how the similarity classes relate to each other in the affine space of $n \times n$ matrices.

The results of Boer, Thijsse, Markus and Parilis [13,57] were extended to matrix pairs by Pokrzywa [60]; he described all possible Kronecker canonical forms of matrix pencils in a neighborhood of a given pencil. In fact, Pokrzywa describes the following partial order on the set $K_{m\times n}$ of Kronecker's canonical forms of $m \times n$ pencils: $A - \lambda B \preccurlyeq C - \lambda D$ if a pencil that is strictly equivalent to $C - \lambda D$ can be obtained by an arbitrarily small perturbation of $A - \lambda B$. The partition of $K_{m\times n}$ into classes of strictly equivalent pencils is a *stratification*, which means that the closure of every class (= stratum) consists of the class itself and a finite union of classes of smaller dimension. Moreover, $A - \lambda B \preccurlyeq C - \lambda D$ if and only if the class of $A - \lambda B$ is contained in the closure of the class of $C - \lambda D$.

Using Pokrzywa's theorem, Bo Kågström and his students [24, 25, 47] develop the software StratiGraph that constructs the Hasse diagram of the poset $K_{m\times n}$, which is also called the closure diagram of the set of $m \times n$ pencils. This diagram shows how the classes of strictly equivalent pencils relate to each other in the affine space of $m \times n$ pencils.

Note that for each matrix pencil $A-\lambda B$, Pokrzywa's theorem describes all Kronecker canonical pencils $C-\lambda D$ such that there exist arbitrarily small E and F for which the Kronecker canonical form of $(A+E)-\lambda(B+F)$ is $C-\lambda D$. Pokrzywa's proof is very abstract and unconstructive; he does not construct $(A+E)-\lambda(B+F)$ explicitly. Even more abstract proof of Pokrzywa's theorem is given by Bongartz in [8]; he uses the representation theory of finite dimensional algebras. The main goal of the thesis is to construct $(A+E)-\lambda(B+F)$ explicitly for each $C-\lambda D$. As a consequence, we get a direct, constructive, and rather elementary proof of Pokrzywa's theorem.

The other goal of the thesis is to give applications of perturbation theory of matrix pairs with respect to equivalence and contragredient equivalence. Two matrix pairs (A,B) and (A',B') of the same size are *contragrediently equivalent* if

$$(A',B')=(S^{-1}AR,R^{-1}BS),$$

in which S,R are nonsingular matrices. Many applications of a contragredient equivalence are given in [45, 63]. If we study a matrix product AB up to similarity transformations, then the matrix pair (A,B) can be considered up to contragredient equivalence since $S^{-1}ABS=S^{-1}ARR^{-1}BS$; see [50].

A canonical form of matrix pairs under contragredient equivalence is given by Dobrovol′skaya and Ponomarev [22]. M.I. García-Planas and V.V. Sergeichuk [39] construct a miniversal deformation of a canonical pair (A,B) for contragredient equivalence; that is, a simple normal form to which all matrix pairs $(A+\widetilde{A},B+\widetilde{B})$ close to (A,B) can be reduced by contragredient equivalence transformations that smoothly depend on the entries of $\widetilde{A}$ and $\widetilde{B}$. Each perturbation $(\widetilde{A},\widetilde{B})$ of (A,B) defines the first order induced perturbation $A\widetilde{B}+\widetilde{A}B$ of the matrix AB, which is the first order summand in the product $(A+\widetilde{A})(B+\widetilde{B})=AB+A\widetilde{B}+\widetilde{A}B+\widetilde{A}\widetilde{B}$.

The following canonical matrix pairs are described in the thesis:

- all canonical matrix pairs (A, B), for which the *first order induced perturbations* $A\widetilde{B} + \widetilde{A}B$ are nonzero for all nonzero perturbations in the normal form of García-Planas and Sergeichuk;
- all canonical matrix pairs (A, B), for which the deformations $(A + \widetilde{A})(B + \widetilde{B}) = AB + A\widetilde{B} + \widetilde{A}B + \widetilde{A}\widetilde{B}$ of AB are versal for all nonzero perturbations in the normal form of García-Planas and Sergeichuk;
- all canonical matrix pairs (A, B) that are structurally stable under nonzero perturbations in the normal form of García-Planas and Sergeichuk. (The concept of structural stability was first introduced by A.A. Andronov and L.S. Pontryagin [2] in 1937 in the qualitative theory of dynamical systems.)

These results are published in [32–34].

P. Van Dooren [70] constructed an algorithm for computing all singular summands of Kronecker's canonical form of a matrix pencil. His algorithm uses only unitary transformations, which improves its numerical stability. In the thesis, Van Dooren's algorithm is extended to square complex matrices up to consimilarity transformations $A \mapsto SA\bar{S}^{-1}$ and to pairs of $m \times n$ complex matrices up to transformations $(A, B) \mapsto (SAR, SB\bar{R})$, in which S and R are nonsingular matrices. This result is publlished in [51].

The thesis is organized as follows:

- An informal introduction into the theory of perturbations of matrices and matrix pairs is given in *Chapter 1*.
- A proof of Pokrzywa's theorem is given in *Chapter 2*.
- Applications of miniversal deformations of matrix pairs with respect to contragredient equivalence are given in *Chapter 3*.

CHAPTER 1

A SURVEY OF KNOWN RESULTS ABOUT PERTURBATIONS OF MATRICES AND MATRIX PENCILS

The purpose of this chapter is to give an informal introduction to the theory of perturbations of matrices and matrix pairs.

All matrices are considered over a field $\mathbb{F} \in \{\mathbb{C}, \mathbb{R}\}$.

1.1. Arnold's miniversal deformations of matrices with respect to similarity

In this section, we formulate Arnold's theorem about miniversal deformations of matrices up to similarity and give a sketch of its constructive proof, because it is described algorithm how to construct the transformation (1.5). Algorithms for constructing this transformation are also discussed in [55].

Since each square matrix is similar to a Jordan matrix, it suffices to study perturbations of Jordan matrices.

For each $A \in \mathbb{C}^{n\times n}$ and a small matrix $X \in \mathbb{C}^{n\times n}$,

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

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 $\{A + XA - AX \,|\, X \in \mathbb{C}^{n\times n}\}$. (By the Lipschitz property [61], the transforming matrix S to a matrix $S^{-1}AS$ near A can be taken in the form $I + X$ with a small X). The vector space

$$T(A) := \{XA - AX \,|\, X \in \mathbb{C}^{n\times n}\} \tag{1.1}$$

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

$$\dim_{\mathbb{C}} T(A), \qquad n^2 - \dim_{\mathbb{C}} T(A) \tag{1.2}$$

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

For each Jordan canonical matrix J whose Jordan blocks are ordered as follows:

$$J = \bigoplus_{i=1}^{t} (J_{m_{i1}}(\lambda_i) \oplus \cdots \oplus J_{m_{ir_i}}(\lambda_i)), \qquad m_{i1} \geqslant m_{i2} \geqslant \ldots \geqslant m_{ir_i} \tag{1.3}$$

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

$$J + \mathcal{D} := \bigoplus_{i=1}^{t} \begin{bmatrix} J_{m_{i1}}(\lambda_i) + 0^{\downarrow} & 0^{\downarrow} & \ldots & 0^{\downarrow} \\ 0^{\leftarrow} & J_{m_{i2}}(\lambda_i) + 0^{\downarrow} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0^{\downarrow} \\ 0^{\leftarrow} & \ldots & 0^{\leftarrow} & J_{m_{ir_i}}(\lambda_i) + 0^{\downarrow} \end{bmatrix} \tag{1.4}$$

in which

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

are blocks whose entries are zeros and stars.

The following theorem was given by Arnold in [3, Theorem 4.4] (or see [4, Section 3.3] or [5, § 30]).

Theorem 1.1 (Arnold [3]). *Let J be the Jordan canonical matrix* (1.3). *Then all matrices $A + X$ that are sufficiently close to A can be simultaneously reduced by some transformation*

$$A + X \mapsto \mathcal{S}(X)^{-1}(A + X)\mathcal{S}(X), \qquad \begin{array}{c} \mathcal{S}(X) \text{ is analytic} \\ \text{at zero, } \mathcal{S}(0) = I \end{array} \tag{1.5}$$

to the form $J + \mathcal{D}$ defined in (1.4)*; its stars represent entries that depend analytically on the entries of X. The number of stars is minimal that can be achieved by transformations of the form* (1.5)*, it is equal to the codimension of the similarity class of J.*

The matrix (1.4) with independent parameters instead of stars is called a *miniversal deformation* of J (see formal definitions in [3] or [4] or [5]).

Remark 1.1.1. The matrix (1.4) is the direct sum of t matrices that are not block triangular. But each Jordan matrix J is permutation similar to the Weyr matrix $J^{\#}$ (all commuting with $J^{\#}$ matrices are upper block triangular). Producing with (1.4) the same transformations of permutation similarity, we obtain an upper block triangular matrix $J^{\#}+\mathcal{D}^{\#}$, which gives a miniversal deformation of $J^{\#}$; see details in [48].

1.2. Simplest miniversal deformations of matrices, matrix pencils, and contragredient matrix pencils

In this section, we provide results of Maria Isabel Garcia Planas and Vladimir V. Sergeichuk (see [39]) about simplest miniversal deformations of matrices, matrix pencils, and contragredient matrix pencils.

All matrices and representations are considered over a field $\mathbb{F} \in \{\mathbb{C}, \mathbb{R}\}$.

1.2.1. Deformations of matrices In the next two sections, we recall how to obtain a simplest miniversal $\mathbb{R}$-deformation of a real matrix and under similarity and a matrix pencil under equivalence and contragredient equivalence using [39].

Let us denote

$$J_r^{\mathbb{C}}(\lambda) = J_r(\lambda) := \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}, \quad J_r := J_r(0); \tag{1.6}$$

and, for $\lambda = a + bi \in \mathbb{C}$ $(b \geqslant 0)$, denote $J_r^{\mathbb{R}}(\lambda) := J_r(\lambda)$ if $b = 0$ and

$$J_r^{\mathbb{R}}(\lambda) := \begin{bmatrix} T_{ab} & I_2 & & \\ & T_{ab} & \ddots & \\ & & \ddots & I_2 \\ & & & T_{ab} \end{bmatrix} \text{ if } b > 0, \text{ where } T_{ab} := \begin{bmatrix} a & b \\ -b & a \end{bmatrix}, \tag{1.7}$$

(the size of $J_r(\lambda)$, $J_r^{\mathbb{C}}(\lambda)$ and $J_r^{\mathbb{R}}(\lambda)$ is $r \times r$).

Clearly, every square matrix over $\mathbb{F} \in \{\mathbb{C}, \mathbb{R}\}$ is similar to a matrix of the form

$$\oplus_i \Phi^{\mathbb{F}}(\lambda_i), \quad \lambda_i \neq \lambda_j \text{ if } i \neq j, \tag{1.8}$$

uniquely determined up to permutations of summands, where

$$\Phi^{\mathbb{F}}(\lambda_i) := \operatorname{diag}(J^{\mathbb{F}}_{s_{i1}}(\lambda_i),\, J^{\mathbb{F}}_{s_{i2}}(\lambda_i), \dots), \quad s_{i1} \geqslant s_{i2} \geqslant \cdots. \tag{1.9}$$

Let

$$\mathcal{H} = [H_{ij}] \tag{1.10}$$

be a parametric block matrix with $p_i \times q_j$ blocks H_{ij} of the form

$$H_{ij} = \begin{bmatrix} * & \\ \vdots & 0 \\ * & \end{bmatrix} \text{ if } p_i \leqslant q_j, \quad H_{ij} = \begin{bmatrix} 0 \\ * \cdots * \end{bmatrix} \text{ if } p_i > q_j, \tag{1.11}$$

where the stars denote independent parameters.

Arnold [3] (see also [5, § 30]) proved that one of the simplest miniversal $\mathbb{C}$-deformations of the matrix (1.8) for $\mathbb{F} = \mathbb{C}$ is $\oplus_i(\Phi^{\mathbb{C}}(\lambda_i) + \mathcal{H}_i)$, where $\mathcal{H}_i$ is of the form (1.10). Galin [29] (see also [5, § 30E]) showed that one of the miniversal $\mathbb{R}$-deformations of the matrix (1.8) for $\mathbb{F} = \mathbb{R}$ is $\oplus_i(\Phi^{\mathbb{R}}(\lambda_i)+\mathcal{H}_{\lambda_i})$, where $\mathcal{H}_\lambda$ ($\lambda \in \mathbb{R}$) is of the form (1.10) and $\mathcal{H}_\lambda$ ($\lambda \notin \mathbb{R}$) is obtained from a matrix of the form (1.10) by the replacement of its entries $\alpha + \beta i$ with 2×2 blocks $T_{\alpha\beta}$ (see (1.7)). For example, a real 4×4 matrix with two Jordan 2×2 blocks with eigenvalues $x \pm iy$ ($y \neq 0$) has a miniversal $\mathbb{R}$-deformation

$$\begin{bmatrix} x & y & 1 & 0 \\ -y & x & 0 & 1 \\ 0 & 0 & x & y \\ 0 & 0 & -y & x \end{bmatrix} + \begin{bmatrix} \alpha_1 & \beta_1 & 0 & 0 \\ -\beta_1 & \alpha_1 & 0 & 0 \\ \alpha_2 & \beta_2 & 0 & 0 \\ -\beta_2 & \alpha_2 & 0 & 0 \end{bmatrix} \tag{1.12}$$

with the parameters $\alpha_1, \beta_1, \alpha_2, \beta_2$. We prove that a simplest miniversal $\mathbb{R}$-deformation of this matrix may be obtained by the replacement of the second column $(\beta_1, \alpha_1, \beta_2, \alpha_2)^T$ in (1.12) with $(0, 0, 0, 0)^T$.

Theorem 1.2 (Arnold [3] for $\mathbb{F} = \mathbb{C}$)**.** *One of the simplest miniversal $\mathbb{F}$-deformations of the canonical matrix (1.8) under similarity over $\mathbb{F} \in \{\mathbb{C}, \mathbb{R}\}$ is $\oplus_i(\Phi^{\mathbb{F}}(\lambda_i) + \mathcal{H}_i)$, where $\mathcal{H}_i$ is of the form (1.10).*

Proof. Let A be the matrix (1.8). We must prove that for every $M \in \mathbb{F}^{m\times m}$ there exists $S \in \mathbb{F}^{m\times m}$ such that

$$M + SA - AS = N, \tag{1.13}$$

where N is obtained from $\oplus_i \mathcal{H}_i$ by replacing its stars with elements of $\mathbb{F}$ and is uniquely determined by M. The matrix A is block-diagonal with diagonal blocks of the form $J_r^{\mathbb{F}}(\lambda)$. We apply the same partition into blocks to M and N and rewrite the equality (1.13) for blocks:

$$M_{ij} + S_{ij}A_j - A_iS_{ij} = N_{ij}.$$

The theorem follows from the next lemma. $\square$

Lemma 1.1. *For given $J_p^{\mathbb{F}}(\lambda)$, $J_q^{\mathbb{F}}(\mu)$, and for every matrix $M \in \mathbb{F}^{p\times q}$ there exists a matrix $S \in \mathbb{F}^{p\times q}$ such that $M + SJ_q^{\mathbb{F}}(\mu) - J_p^{\mathbb{F}}(\lambda)S = 0$ if $\lambda \neq \mu$, and $M + SJ_q^{\mathbb{F}}(\mu) - J_p^{\mathbb{F}}(\lambda)S = H$ if $\lambda = \mu$, where H is of the form (1.11) with elements from $\mathbb{F}$ instead of the stars; moreover, H is uniquely determined by M.*

1.2.2. Deformations of matrix pencils The canonical form problem for pairs of matrices $A, B \in \mathbb{F}^{m\times n}$ under transformations of simultaneous equivalence

$$(A, B) \mapsto (SAR^{-1}, SBR^{-1}), \quad S \in \mathrm{GL}(m, \mathbb{F}), \quad R \in \mathrm{GL}(n, \mathbb{F}),$$

(that is, for representations of the quiver $\bullet \rightrightarrows \bullet$) was solved by Kronecker: each pair is uniquely, up to permutation of summands, reduced to a direct sum of pairs of the form (see (1.6)–(1.7))

$$(I, J_r^{\mathbb{F}}(\lambda)),\ (J_r, I),\ (L_r, R_r),\ (L_r^T, R_r^T), \tag{1.14}$$

where $\lambda = a + bi \in \mathbb{C}$ $(b \geqslant 0$ if $\mathbb{F} = \mathbb{R})$ and

$$L_r = \begin{bmatrix} 1 & & 0 \\ 0 & \ddots & \\ & \ddots & 1 \\ 0 & & 0 \end{bmatrix}, \quad R_r = \begin{bmatrix} 0 & & 0 \\ 1 & \ddots & \\ & \ddots & 0 \\ 0 & & 1 \end{bmatrix} \tag{1.15}$$

are matrices of size $r \times (r-1)$, $r \times (r-1)$, $r \geqslant 1$.

A miniversal, but not a simplest miniversal, deformation of the canonical pairs of matrices under simultaneous similarity was obtained in [23], partial cases were considered in [6]– [36].

Denote by $0^{\uparrow}$ (resp., $0^{\downarrow}$, $0^{\leftarrow}$, $0^{\rightarrow}$) a matrix, in which all entries are zero except for the entries of the first row (resp., the last row, the first column, the last column) that are independent parameters; and denote by Z the $p \times q$ matrix, in which the first $\max\{q-p, 0\}$ entries of the first row are independent parameters and the other entries are zeros:

$$0^{\uparrow} = \begin{bmatrix} * & \cdots & * \\ 0 & \cdots & 0 \\ \hdotsfor{3} \\ 0 & \cdots & 0 \end{bmatrix}, \quad Z = \begin{bmatrix} * & \cdots & * & 0 & \cdots & 0 \\ & & & & \ddots & \\ & 0 & & 0 & \cdots & 0 \end{bmatrix}. \tag{1.16}$$

We arrange the Jordan blocks in a Jordan matrix with a single eigenvalue as follows:

$$J_{k_1,\dots,k_s}(\lambda) := J_{k_1}(\lambda) \oplus \cdots \oplus J_{k_s}(\lambda), \qquad k_1 \leqslant k_2 \leqslant \cdots \leqslant k_s.$$

Define the matrix

$$\widetilde{J}_{k_1,\dots,k_s}(\lambda) := \begin{bmatrix} J_{k_1}(\lambda) + 0^{\leftarrow} & 0^{\leftarrow} & \dots & 0^{\leftarrow} \\ 0^{\downarrow} & J_{k_2}(\lambda) + 0^{\leftarrow} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0^{\leftarrow} \\ 0^{0^{\downarrow}} & \dots & 0^{\downarrow} & J_{k_s}(\lambda) + 0^{\leftarrow} \end{bmatrix}. \tag{1.17}$$

Edelman, Elmroth, and Kågström [23] construct miniversal deformations of matrix pairs under equivalence; however, their deformations contain repeating parameters. We use simpler miniversal deformations that are constructed by Garcia-Planas and Sergeichuk [39]; all parameters in their deformations are independent.

Theorem 1.3 ([39]). *Let $\mathcal{A}$ be the Kronecker pair* (1.14)*. Then all matrix pairs $\mathcal{A}+\mathcal{X}$ that are sufficiently close to $\mathcal{A}$ can be simultaneously reduced by some equivalence transformation*

$$\mathcal{A}+\mathcal{X} \mapsto R(\mathcal{X})^{-1}(\mathcal{A}+\mathcal{X})S(\mathcal{X}), \qquad \begin{array}{l}\textit{matrices } R(\mathcal{X}) \textit{ and } S(\mathcal{X}) \textit{ are}\\ \textit{analytic and identity at } (0,0)\end{array} \tag{1.18}$$

to the form

$$\left(\left[\begin{array}{cccc|c|c|ccc}
L^T_{m_1} & & 0 & & & 0^{\downarrow} & & & \\
& L^T_{m_2} & & & 0 & 0^{\downarrow} & 0^{\rightarrow} & \dots & 0^{\rightarrow}\\
& & \ddots & & & \vdots & & & \\
& & & L^T_{m_{\underline{s}}} & & 0^{\downarrow} & & & \\
\cline{5-9}
& & & & I & 0 & & 0 & \\
\cline{5-9}
& & & & & \widetilde{J}_0 & 0^{\rightarrow} & \dots & 0^{\rightarrow}\\
\cline{6-9}
& & & & & & L_{n_1} & & 0\\
& & & & & & & L_{n_2} & \\
& 0 & & & & & & & \ddots\\
& & & & & & & & L_{n_{\overline{s}}}
\end{array}\right],
\left[\begin{array}{cccc|c|c|cccc}
R^T_{m_1} & Z & \dots & Z & 0^{\uparrow} & & & & & 0^{\uparrow}\\
& R^T_{m_2} & \ddots & \vdots & 0^{\uparrow} & 0 & & & & 0^{\uparrow}\\
& & \ddots & Z & \vdots & & & & & \vdots\\
& & & R^T_{m_{\underline{s}}} & 0^{\uparrow} & & & & & 0^{\uparrow}\\
\cline{5-10}
& & & & \widetilde{J} & 0 & 0^{\leftarrow} & & \dots & 0^{\leftarrow}\\
\cline{5-10}
& & & & & I & & & 0 & \\
\cline{6-10}
& & & & & & R_{n_1} & Z^T & \dots & Z^T\\
& & & & & & & R_{n_2} & \ddots & \vdots\\
& 0 & & & & & & & \ddots & Z^T\\
& & & & & & & & & R_{n_{\overline{s}}}
\end{array}\right]\right) \tag{1.19}$$

in which

$$\widetilde{J} := \bigoplus_{i=1}^{t-1} \widetilde{J}_{k_{i1},\dots,k_{is_i}}(\lambda_i), \qquad \widetilde{J}_0 := \widetilde{J}_{k_{t1},\dots,k_{ts_t}}(0)$$

(see (1.16)*) and the stars are replaced by complex numbers that depend analytically on the entries of the pair $\mathcal{X}$. The number of stars is minimal that can be achieved by equivalence transformations of the form* (1.18)*; this number is equal to the codimension of $\langle\mathcal{A}\rangle$.*

By a *miniversal normal pair* we mean a matrix pair that is obtained from (1.19) by replacing its stars by complex numbers.

1.2.3. Deformations of contragredient matrix pencils The canonical form problem for pairs of matrices $A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times m}$ under transformations of contragredient equivalence

$$(A,B) \mapsto (SAR^{-1}, RBS^{-1}), \quad S \in \mathrm{GL}(m,\mathbb{F}), \quad R \in \mathrm{GL}(n,\mathbb{F}),$$

(i.e., for representations of the quiver $\bullet \rightleftarrows \bullet$) was solved in [22,45]: each pair is uniquely, up to permutation of cells $J_r^{\mathbb{F}}(\lambda)$ in $\oplus_i \Phi^{\mathbb{F}}(\lambda_i)$, reduced to a direct sum

$$\bigoplus_j (I,C) \oplus \bigoplus_{j=1}^{t_1} (I_{r_{1j}}, J_{r_{1j}}) \oplus \bigoplus_{j=1}^{t_2} (J_{r_{2j}}, I_{r_{2j}}) \oplus \bigoplus_{j=1}^{t_3} (F_{r_{3j}}, G_{r_{3j}}) \oplus \bigoplus_{j=1}^{t_4} (G_{r_{4j}}, F_{r_{4j}}) \quad (1.20)$$

where C is of the form (1.8), $r_{i1} \geqslant r_{i2} \geqslant \ldots \geqslant r_{it_i}$, and

$$J_r(\lambda) := \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix} (\lambda \in \mathbb{C}),$$

$$F_r := \begin{bmatrix} 1 & & 0 \\ 0 & \ddots & \\ & \ddots & 1 \\ 0 & & 0 \end{bmatrix}, \qquad G_r := \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & 0 & 1 \end{bmatrix} \quad (1.21)$$

are $r\times r$, $r\times(r-1)$, $(r-1)\times r$ matrices; we denote by 0_{mn} the zero matrix of size $m\times n$, where $m,n \in \{0,1,2,\ldots\}$.

Theorem 1.4. *One of the simplest miniversal $\mathbb{F}$-deformations of the canonical pair (1.20) under contragredient equivalence over $\mathbb{F} \in \{\mathbb{C},\mathbb{R}\}$ is the direct sum of $(I,\tilde{C})$ ($\tilde{C}$ is a simplest miniversal $\mathbb{F}$-deformation of C*

under similarity, see Theorem 1.2) and

$$\left(\left[\begin{array}{c|c|c} \oplus_j I_{r_{1j}} & 0 & 0 \\ \hline 0 & \oplus_j J_{r_{2j}} + \mathcal{H} & \mathcal{H} \\ \hline 0 & \mathcal{H} & \begin{matrix} P_3 & \mathcal{H} \\ 0 & Q_4 \end{matrix} \end{array}\right], \left[\begin{array}{c|c|c} \oplus_j J_{r_{1j}} + \mathcal{H} & \mathcal{H} & \mathcal{H} \\ \hline \mathcal{H} & \oplus_j I_{r_{2j}} & 0 \\ \hline \mathcal{H} & 0 & \begin{matrix} Q_3 & 0 \\ \mathcal{H} & P_4 \end{matrix} \end{array}\right]\right),$$

where

$$P_l = \begin{bmatrix} F_{r_{l1}} + H & H & \cdots & H \\ & F_{r_{l2}} + H & \ddots & \vdots \\ & & \ddots & H \\ 0 & & & F_{r_{lt_l}} + H \end{bmatrix}, \ Q_l = \begin{bmatrix} G_{r_{l1}} & & & 0 \\ H & G_{r_{l2}} & & \\ \vdots & \ddots & \ddots & \\ H & \cdots & H & G_{r_{lt_l}} \end{bmatrix}$$

$(l = 3, 4)$, *$\mathcal{H}$ and H are matrices of the form (1.10) and (1.11), the stars denote independent parameters.*

1.3. Pokrzywa's theorem for matrix pencils

Pokrzywa [60] described the following partial order on the set $K_{m\times n}$ of Kronecker's canonical forms of $m \times n$ pencils:

$$A - \lambda B \leqslant C - \lambda D$$

if and only if a pencil that is strictly equivalent to $C - \lambda D$ can be obtained by an arbitrarily small perturbation of $A - \lambda B$.

The partition of $K_{m\times n}$ into classes of strictly equivalent pencils is a *stratification*, which means that the closure of every class (= stratum) consists of the class itself and a finite union of classes of smaller dimension. Moreover, $A - \lambda B \leqslant C - \lambda D$ if and only if the class of $A - \lambda B$ is contained in the closure of the class of $C - \lambda D$.

Let us recall Pokrzywa's theorem, which describes all possible Kronecker's canonical forms of matrix pencils that are arbitrarily close to a given pencil. We use notation from [67].

The orbit $\mathcal{O}(\mathcal{M})$ of an $m \times n$ matrix pencil $\mathcal{M}(\lambda) = A - \lambda B$ is the set of matrix pencils strictly equivalent to $\mathcal{M}(\lambda)$:

$$\mathcal{O}(\mathcal{M}) = \{P\mathcal{M}(\lambda)Q : P \in \mathbb{C}^{m\times m}, Q \in \mathbb{C}^{n\times n}, \quad P, Q \quad \text{are nonsingular}\}$$

These orbits are manifolds in the vector space $\mathbb{C}^{2mn}$, and we refer to the *codimension* of $\mathcal{O}(\mathcal{M})$ as the codimension in this space. We denote by $\overline{\mathcal{O}(\mathcal{M})}$ the closure of this orbit.

The most significant element of the orbit $\mathcal{O}(\mathcal{M})$ is the Kronecker canonical form (e.g., see [30]) of $\mathcal{M}(\lambda)$. The Kronecker canonical form is the direct sum of the right singular, left singular, and regular structures, consisting of L_k blocks of dimension $k \times (k + 1)$ for the *right singular structure* and L_k^T blocks for the *left singular structure*. The *regular structure* consists of Jordan blocks $J_k(\mu)$ corresponding to eigenvalue μ, and N_k corresponding to the infinite eigenvalue. The Kronecker canonical form of $\mathcal{M}(\lambda)$ determines uniquely the orbit $\mathcal{O}(\mathcal{M})$, and, in particular, it fully determines the codimension of $\mathcal{O}(\mathcal{M})$ [16, Theorem 2.2].

The *dominance ordering* in the set of sequences of nonnegative integers specifies that $(a_1, a_2, \dots) \geqslant (b_1, b_2, \dots)$ if $a_1 + a_2 + \dots + a_i \geqslant b_1 + b_2 + \dots + b_i$ for $i = 1, 2, \dots$. We say that $(a_1, a_2, \dots) > (b_1, b_2, \dots)$ if $(a_1, a_2, \dots) \geqslant (b_1, b_2, \dots)$ and $(a_1, a_2, \dots) \neq (b_1, b_2, \dots)$ [24, Section 2.1].

For every matrix pencil $\mathcal{M}(\lambda)$ with rank r, which is defined in [30, Chapter VI] as the order of its largest minor that is not equal to the zero polynomial in λ, we consider the following sequences defined in [24]:

$$\mathcal{R}(\mathcal{M}) + r = (r_0 + r, r_1 + r, r_2 + r, \dots);$$

where r_i is the number of right singular blocks in the Kronecker canonical form of $\mathcal{M}(\lambda)$ with $j \geqslant i$;

$$\mathcal{L}(\mathcal{M}) + r = (l_0 + r, l_1 + r, l_2 + r, \dots);$$

where l_i is the number of left singular blocks in the Kronecker canonical form of $\mathcal{M}(\lambda)$ with $j \geqslant i$; and for every $\mu \in \mathbb{C} \cup \{\infty\}$,

$$\mathcal{J}_\mu(\mathcal{M}) + p = (\omega_1(\mu) + p, \omega_2(\mu) + p, \dots);$$

where $\omega_i(\mu)$ is the number of Jordan blocks associated with the eigenvalue μ of dimension greater than or equal to i in the regular structure of the Kronecker canonical form of $\mathcal{M}(\lambda)$, and p is the number of right singular blocks in the Kronecker canonical form of $\mathcal{M}(\lambda)$. These sequences allow us to describe the inclusion relationships between the closures of the orbits of two matrix pencils. The corresponding theorem is obtained in [60], and later reformulated in [24] and [67]. We state it as in [67, theorem 2.1].

Theorem 1.5 (see [24, Theorem 3.1] and [60, Theorem 3])**.** *Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two $m \times n$ complex matrix pencils with $p(\mathcal{M}_1)$ and $p(\mathcal{M}_2)$ right singular blocks in their Kronecker canonical forms, respectively. Then $\overline{\mathcal{O}(\mathcal{M})} \supseteq \overline{\mathcal{O}(\mathcal{M})}$ if and only if the following relations hold:*

(i) $\mathcal{R}(\mathcal{M}_1) + \operatorname{rank}(\mathcal{M}_1) \geqslant \mathcal{R}(\mathcal{M}_2) + \operatorname{rank}(\mathcal{M}_2)$,

(ii) $\mathcal{L}(\mathcal{M}_1) + \operatorname{rank}(\mathcal{M}_1) \geqslant \mathcal{L}(\mathcal{M}_2) + \operatorname{rank}(\mathcal{M}_2)$,

(iii) $\mathcal{J}_\mu(\mathcal{M}_1) + p(\mathcal{M}_1) \leqslant \mathcal{J}_\mu(\mathcal{M}_2) + p(\mathcal{M}_2)$

for all $\mu \in \mathbb{C} \cup \{\infty\}$.

Pokrzywa's proof is very abstract and unconstructive. Even more abstract proof of Pokrzywa's theorem was given by Bongartz in [8]; he uses the representation theory of finite dimensional algebras. We prove Pokrzywa's theorem in Boley's form [10] (see also Dopico and Dmytryshyn [19]).

CHAPTER 2

STRATIFICATION THEORY OF MATRIX PAIRS UNDER EQUIVALENCE AND CONTRAGREDIENT EQUIVALENCE

2.1. Main results

All matrices in this section are complex matrices and both matrices in each matrix pair have the same size; we call it the *size of the pair*. Two matrix pairs (A, A') and (B, B') are *equivalent*, we write $(A, A') \sim (B, B')$, if there exist nonsingular matrices R and S such that

$$R(A, A')S := (RAS, RA'S) = (B, B').$$

The *orbit* of $\mathcal{A} = (A, A')$ is the set $\langle\mathcal{A}\rangle$ of all pairs that are equivalent to $\mathcal{A}$. Define the partial ordering on the set of orbits of pairs of the same size as follows: $\langle\mathcal{A}\rangle \leqslant \langle\mathcal{B}\rangle$ if $\langle\mathcal{A}\rangle$ is contained in the closure of $\langle\mathcal{B}\rangle$. This means that in each neighborhood of $\mathcal{A}$ there is a pair that is equivalent to $\mathcal{B}$.

For each positive integer n, we define the matrices

$$L_n := \begin{bmatrix} 1 & 0 & & 0 \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{bmatrix}, \quad R_n := \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & 0 & 1 \end{bmatrix} \quad ((n-1)\text{-by-}n),$$

$$J_n(\lambda) := \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix} \quad (n\text{-by-}n,\ \lambda \in \mathbb{C}).$$

We also define the matrices

$$0^{\nwarrow} := \begin{bmatrix} 1\,0\dots 0 \\ 0 \end{bmatrix},\quad 0^{\nearrow} := \begin{bmatrix} 0\dots 0\,1 \\ 0 \end{bmatrix},\quad 0^{\swarrow} := \begin{bmatrix} 0 \\ 1\,0\dots 0 \end{bmatrix},\quad 0^{\searrow} := \begin{bmatrix} 0 \\ 0\dots 0\,1 \end{bmatrix},$$

whose sizes will be clear from the context.

The matrix pairs

$$\mathcal{L}_n := (L_n, R_n), \qquad \mathcal{L}_n^T := (L_n^T, R_n^T), \tag{2.1}$$

$$\mathcal{D}_n(\lambda) := \begin{cases} (I_n, J_n(\lambda)) & \text{if } \lambda \in \mathbb{C}, \\ (J_n(0), I_n) & \text{if } \lambda = \infty. \end{cases} \tag{2.2}$$

are called *indecomposable Kronecker pairs*. Their direct sums are called *Kronecker pairs*. By Kronecker's theorem, each matrix pair $\mathcal{A}$ is equivalent to a Kronecker pair, which is called the *Kronecker canonical form* of $\mathcal{A}$, and which is determined by $\mathcal{A}$ uniquely up to permutations of direct summands.

Theorem I. Let $\mathcal{A}$ and $\mathcal{B}$ be nonequivalent Kronecker pairs. Then $\langle\mathcal{A}\rangle < \langle\mathcal{B}\rangle$ if and only if $\mathcal{B}$ can be obtained from $\mathcal{A}$ by permutations of direct summands and replacements of pairs of summands of the types (i)–(vi) listed below, in which $m, n \in \{1, 2, \dots\}$, $\lambda \in \mathbb{C} \cup \infty$. The notation $\mathcal{P} \downarrow \mathcal{Q}$ means that $\mathcal{P}$ is replaced by $\mathcal{Q}$.

(i) $\mathcal{L}_m \oplus \mathcal{L}_n \ \downarrow\ \mathcal{L}_{m+1} \oplus \mathcal{L}_{n-1}$ in which $m + 2 \leqslant n$.;

(ii) $\mathcal{L}_m^T \oplus \mathcal{L}_n^T \ \downarrow\ \mathcal{L}_{m+1}^T \oplus \mathcal{L}_{n-1}^T$ in which $m + 2 \leqslant n$.

(iii) $\mathcal{L}_m \oplus \mathcal{D}_n(\lambda) \ \downarrow\ \mathcal{L}_{m+1} \oplus \mathcal{D}_{n-1}(\lambda)$ (the summands $\mathcal{D}_0(\lambda)$ are omitted).

(iv) $\mathcal{L}_m^T \oplus \mathcal{D}_n(\lambda) \ \downarrow\ \mathcal{L}_{m+1}^T \oplus \mathcal{D}_{n-1}(\lambda)$.

(v) $\mathcal{D}_m(\lambda) \oplus \mathcal{D}_n(\lambda) \ \downarrow\ \mathcal{D}_{m-1}(\lambda) \oplus \mathcal{D}_{n+1}(\lambda)$ in which $m \leqslant n$.

(vi) $\mathcal{L}_m^T \oplus \mathcal{L}_n \ \downarrow\ \mathcal{D}_{r_1}(\mu_1) \oplus \dots \oplus \mathcal{D}_{r_k}(\mu_k)$, in which $\mu_1, \dots, \mu_k \in \mathbb{C} \cup \infty$ are distinct and $r_1 + \dots + r_k = m + n - 1$.

Up to permutation of summands, each Kronecker pair is of the form

$$\begin{gathered}\mathcal{A} := \bigoplus_{i=1}^{\underline{s}} \mathcal{L}_{m_i}^T \oplus \bigoplus_{i=1}^{\overline{s}} \mathcal{L}_{n_i} \oplus \bigoplus_{i=1}^{t} \Big(\mathcal{D}_{k_{i1}}(\lambda_i) \oplus \dots \oplus \mathcal{D}_{k_{is_i}}(\lambda_i)\Big),\\ m_1 \leqslant \dots \leqslant m_{\underline{s}}, \quad n_1 \leqslant \dots \leqslant n_{\overline{s}}, \quad k_{i1} \leqslant \dots \leqslant k_{is_i}\ (i = 1, \dots, t),\end{gathered} \tag{2.3}$$

in which $\lambda_1, \dots, \lambda_t \in \mathbb{C} \cup \infty$ are distinct. The numbers $\underline{s}, s_1, \dots, s_t, \overline{s}$ can be zero, which means that the corresponding direct summands in (2.6) are absent.

An orbit $\langle \mathcal{B} \rangle$ *immediately succeeds* $\langle \mathcal{A} \rangle$ if $\langle \mathcal{A} \rangle < \langle \mathcal{B} \rangle$ and there exists no $\langle \mathcal{C} \rangle$ such that $\langle \mathcal{A} \rangle < \langle \mathcal{C} \rangle < \langle \mathcal{B} \rangle$.

Theorem II. Let $\mathcal{A}$ be the Kronecker pair (2.6). An orbit $\mathcal{O}$ immediately succeeds the orbit $\langle \mathcal{A} \rangle$ if and only if $\mathcal{O}$ is the orbit of a pair that if obtained from $\mathcal{A}$ by one of the replacements

(i′) $\mathcal{L}_{n_i} \oplus \mathcal{L}_{n_j} \ \downarrow\ \mathcal{L}_{n_i+1} \oplus \mathcal{L}_{n_j-1}$, in which either $j = i+1$ and $n_i + 2 \leqslant n_{i+1}$, or $j = i+2$ and $n_i + 1 = n_{i+1} = n_{i+2} - 1$,

(ii′) $\mathcal{L}_{m_i}^T \oplus \mathcal{L}_{m_j}^T \ \downarrow\ \mathcal{L}_{m_i+1}^T \oplus \mathcal{L}_{m_j-1}^T$, in which either $j = i+1$ and $m_i + 2 \leqslant m_{i+1}$, or $j = i+2$ and $m_i + 1 = m_{i+1} = m_{i+2} - 1$,

(iii′) $\mathcal{L}_{n_{\overline{s}}} \oplus \mathcal{D}_{k_{is_i}}(\lambda_i) \ \downarrow\ \mathcal{L}_{n_{\overline{s}}+1} \oplus \mathcal{D}_{k_{is_i}-1}(\lambda_i)$,

(iv′) $\mathcal{L}_{m_{\underline{s}}}^T \oplus \mathcal{D}_{k_{is_i}}(\lambda_i) \ \downarrow\ \mathcal{L}_{m_{\underline{s}}+1}^T \oplus \mathcal{D}_{k_{is_i}-1}(\lambda_i)$,

(v′) $\mathcal{D}_{k_{ij}}(\lambda_i) \oplus \mathcal{D}_{k_{i,j+1}}(\lambda_i) \ \downarrow\ \mathcal{D}_{k_{ij}-1}(\lambda_i) \oplus \mathcal{D}_{k_{i,j+1}+1}(\lambda_i)$,

(vi′) $\mathcal{L}_{n_{\overline{s}}} \oplus \mathcal{L}_{m_{\underline{s}}}^T \ \downarrow\ \mathcal{D}_{r_1}(\mu_1) \oplus \dots \oplus \mathcal{D}_{r_q}(\mu_q)$, in which $q \geqslant t$,

$$\mu_1 = \lambda_1, \ \dots, \ \mu_t = \lambda_t, \qquad k_{1s_1} \leqslant r_1, \ \dots, \ k_{ts_t} \leqslant r_t, \tag{2.4}$$

$\mu_1, \dots, \mu_q \in \mathbb{C} \cup \infty$ are distinct, and $r_1 + \dots + r_q = n_{\overline{s}} + m_{\underline{s}} - 1$,

which are special cases of the replacements (i)–(vi) from the first main theorem.

Define the matrices

$$\Delta_r(\varepsilon) := \begin{bmatrix} 0\dots0\,\varepsilon\,0\dots0 \\ 0 \end{bmatrix}, \qquad \nabla_r(\varepsilon) := \begin{bmatrix} 0 \\ 0\dots0\,\varepsilon\,0\dots0 \end{bmatrix} \tag{2.5}$$

(whose sizes will be clear from the context), in which ε is an arbitrary nonzero complex number that is located in the rth column. We often write Δ_r and ∇_r omitting ε. Set $\Delta_0 = \nabla_0 := 0$.

For each orbit $\langle\mathcal{A}\rangle$, its *lower cone* is the set $\langle\mathcal{A}\rangle^{\vee}$ of all orbits $\langle\mathcal{B}\rangle$ such that $\langle\mathcal{A}\rangle \leqslant \langle\mathcal{B}\rangle$. In the course of the proof of Theorem I, we describe the lower cones of all direct sums of two indecomposable Kronecker pairs.

Let us rearrange the direct summands of (2.3) as follows:

$$\begin{aligned} \mathcal{A} :=& \mathcal{L}^T_{m_1} \oplus \mathcal{L}^T_{m_2} \oplus \dots \oplus \mathcal{L}^T_{m_{\underline{s}}} \\ & \oplus \bigoplus_{i=1}^{t} \Big(\mathcal{D}_{k_{i1}}(\lambda_i) \oplus \mathcal{D}_{k_{i2}}(\lambda_i) \oplus \dots \oplus \mathcal{D}_{k_{is_i}}(\lambda_i)\Big) \\ & \oplus \mathcal{L}_{n_{\overline{s}}} \oplus \mathcal{L}_{n_{\overline{s}-1}} \oplus \dots \oplus \mathcal{L}_{n_1}, \end{aligned} \tag{2.6}$$

$$m_1 \leqslant \dots \leqslant m_{\underline{s}}, \quad k_{i1} \leqslant \dots \leqslant k_{is_i} \ (i = 1, \dots, t), \quad n_1 \leqslant \dots \leqslant n_{\overline{s}}.$$

It follows from Theorem I, that each immediate successor of $\langle\mathcal{A}\rangle$ is the orbit of a pair that is obtained by a perturbation of only one pair of the upper diagonal blocks of $\mathcal{A}$.

Corollary 2.1 (of Theorem I)**.** *Let* $\mathcal{A} = ([A_{ij}], [A'_{ij}])$ *be the Kronecker pair* (2.6) *partitioned such that the pairs of diagonal blocks* $(A_{11}, A'_{11}), (A_{22}, A'_{22}), \dots$ *are the direct summands* $\mathcal{L}^T_{m_1}, \mathcal{L}^T_{m_2}, \dots, \mathcal{L}^T_{m_{\underline{s}}}$, $\mathcal{D}_{k_{11}}(\lambda_1), \mathcal{D}_{k_{12}}(\lambda_1), \dots$ *from* (2.6)*. Then each immediate successor of* $\langle\mathcal{A}\rangle$ *is the orbit of some matrix pair that is obtained from* $\mathcal{A}$ *by an arbitrarily small perturbation of only one pair* (A_{ij}, A'_{ij}) *with* $i < j$ *of its upper diagonal blocks.*

We move backwards in the next sections: we first give an independent proof of a weaker form of Theorem IV in Corollary 2.3, and then we prove Theorem I in the next sections.

2.2. Theorem II follows from Theorem I

In this section, we derive Theorem II from Theorem I. It is sufficient to prove the following statement:

Let a Kronecker pair $\mathcal{B}$ be obtained from a Kronecker pair $\mathcal{A}$ by a replacement (j) $\in \{(\mathrm{i}), (\mathrm{ii}), \dots, (\mathrm{vi})\}$ *from Theorem I. Then $\langle\mathcal{B}\rangle$ immediately succeeds $\langle\mathcal{A}\rangle$ if and only if* (j) *is the replacement* (j′) *from Theorem II.* (2.7)

Let us prove this statement for the pair $\mathcal{A}$ given in (2.6).

Case 1: (j) *is the replacement* (i)*:*

$$\mathcal{L}_{n_i} \oplus \mathcal{L}_{n_j} \ \downarrow\ \mathcal{L}_{n_i+1} \oplus \mathcal{L}_{n_j-1}, \quad \text{in which } n_i + 2 \leqslant n_j. \tag{2.8}$$

$\Longrightarrow$. Let $\langle\mathcal{B}\rangle$ immediately succeed $\langle\mathcal{A}\rangle$. We must prove that (2.8) is the replacement (i′). To the contrary, let $i+2 \leqslant j$, $n_i < n_{i+1} < n_j$, and $n_i+3 \leqslant n_j$. If $n_i + 2 \leqslant n_{i+1}$, then (2.8) is the following composition of replacements of type (i):

$$\mathcal{L}_{n_i} \oplus \mathcal{L}_{n_{i+1}} \oplus \mathcal{L}_{n_j} \ \downarrow\ \mathcal{L}_{n_i+1} \oplus \mathcal{L}_{n_{i+1}-1} \oplus \mathcal{L}_{n_j} \ \downarrow\ \mathcal{L}_{n_i+1} \oplus \mathcal{L}_{n_{i+1}} \oplus \mathcal{L}_{n_j-1}.$$

By Theorem I,

$$\langle \mathcal{L}_{n_i} \oplus \mathcal{L}_{n_{i+1}} \oplus \mathcal{L}_{n_j} \rangle < \langle \mathcal{L}_{n_i+1} \oplus \mathcal{L}_{n_{i+1}-1} \oplus \mathcal{L}_{n_j} \rangle < \langle \mathcal{L}_{n_i+1} \oplus \mathcal{L}_{n_{i+1}} \oplus \mathcal{L}_{n_j-1} \rangle,$$

and so $\langle\mathcal{B}\rangle$ is not an immediate successor of $\langle\mathcal{A}\rangle$. If $n_i + 1 = n_{i+1}$, then $n_{i+1} + 2 \leqslant n_j$ and (2.8) is the following composition of replacements of type (i):

$$\mathcal{L}_{n_i} \oplus \mathcal{L}_{n_{i+1}} \oplus \mathcal{L}_{n_j} \ \downarrow\ \mathcal{L}_{n_i} \oplus \mathcal{L}_{n_{i+1}+1} \oplus \mathcal{L}_{n_j-1} \ \downarrow\ \mathcal{L}_{n_i+1} \oplus \mathcal{L}_{n_{i+1}} \oplus \mathcal{L}_{n_j-1}.$$

Thus, $\langle\mathcal{B}\rangle$ is not an immediate successor of $\langle\mathcal{A}\rangle$ too.

$\Longleftarrow$. Let $\mathcal{B}$ be obtained from $\mathcal{A}$ by replacement (i′). Let $\mathcal{B}$ be also obtained from $\mathcal{A}$ by a sequence

$$\mathcal{A} = \mathcal{A}_1 \xmapsto{\varphi_1} \mathcal{A}_2 \xmapsto{\varphi_2} \mathcal{A}_3 \xmapsto{\varphi_3} \cdots \xmapsto{\varphi_p} \mathcal{A}_{p+1} = \mathcal{B} \tag{2.9}$$

of replacements of types (i)–(vi). In order to show that $\langle\mathcal{B}\rangle$ is an immediate successor of $\langle\mathcal{A}\rangle$, we must prove that $p = 1$.

All replacements $\varphi_1, \dots, \varphi_p$ are not of

- type (vi) since $\mathcal{A}$ and $\mathcal{B}$ have the same number $\underline{s}$ of summands $\mathcal{L}_i^T$, but (vi) decreases the number $\underline{s}$ and this number cannot be restored by (i)–(v);

- type (iv) since it increases the number $m_1 + \dots + m_{\underline{s}}$ whereas this number is not changed by (i), (ii), (iii), and (v);

- type (iii) since it increases $n_1 + \dots + n_{\tilde{s}}$;

- type (v) with $\lambda = \lambda_i$ since it increases $\sum_{p<q}(k_{iq} - k_{ip})$ whereas this number is not changed by (i) and (ii);

- type (ii) since it decreases $\sum_{i<j}(m_j - m_i)$.

Therefore, all $\varphi_1, \dots, \varphi_p$ are replacements of type (i). Since each replacement (i′) is not the composition of several replacements of type (i), $p = 1$, and so $\langle\mathcal{B}\rangle$ is an immediate successor of $\langle\mathcal{A}\rangle$. We have proved (2.7) in Case 1.

Case 2: (j) *is the replacement* (ii). The statement (2.7) is proved in this case by transposing the matrices in Case 1.

Case 3: (j) *is the replacement* (iii)*:*

$$\mathcal{L}_n \oplus \mathcal{D}_k(\lambda_i) \;\downarrow\; \mathcal{L}_{n+1} \oplus \mathcal{D}_{k-1}(\lambda_i), \tag{2.10}$$

in which $(n, k) = (n_l, k_{ij})$ for some l and j.

$\Longrightarrow$. To the contrary, suppose that (2.10) is not (iii′); that is, $n < n_{\overline{s}}$ or $k < k_{is_i}$. If $n < n_{\overline{s}}$, then (2.10) is the composition of replacements of types (i) and (iii):

$$\mathcal{L}_n \oplus \mathcal{L}_{n_{\overline{s}}} \oplus \mathcal{D}_k(\lambda_i) \;\downarrow\; \mathcal{L}_n \oplus \mathcal{L}_{n_{\overline{s}}+1} \oplus \mathcal{D}_{k-1}(\lambda_i) \;\downarrow\; \mathcal{L}_{n+1} \oplus \mathcal{L}_{n_{\overline{s}}} \oplus \mathcal{D}_{k-1}(\lambda_i).$$

If $k < k_{is_i}$, then

$$\mathcal{L}_n \oplus \mathcal{D}_k(\lambda_i) \oplus \mathcal{D}_{k_{is_i}}(\lambda_i) \downarrow \mathcal{L}_{n+1} \oplus \mathcal{D}_k(\lambda_i) \oplus \mathcal{D}_{k_{is_i}-1}(\lambda_i) \downarrow \mathcal{L}_{n+1} \oplus \mathcal{D}_{k-1}(\lambda_i) \oplus \mathcal{D}_{k_{is_i}}(\lambda_i)$$

By the first main theorem, $\langle \mathcal{B} \rangle$ is not an immediate successor of $\langle \mathcal{A} \rangle$.

$\Longleftarrow$. Let $\mathcal{B}$ be obtained from $\mathcal{A}$ by replacement (iii′). Let $\mathcal{B}$ be also obtained from $\mathcal{A}$ by a sequence $\mathcal{A} = \mathcal{A}_1 \xmapsto{\varphi_1} \mathcal{A}_2 \xmapsto{\varphi_2} \cdots \xmapsto{\varphi_p} \mathcal{A}_{p+1} = \mathcal{B}$ of replacements of types (i)–(vi).

All replacements $\varphi_1, \dots, \varphi_p$ are not of

- type (vi) since it decreases the number $\underline{s}$;
- type (i) since it increases lexicographically $(n_1, n_2, \dots, n_{\underline{s}})$;
- types (ii) and (iv) since they change the sequence $(m_1, m_2, \dots, m_{\overline{s}})$;
- type (v) with $\lambda = \lambda_l$ since it decreases lexicographically $(k_{i1}, k_{i2}, \dots, k_{is_i})$.

Therefore, all $\varphi_1, \dots, \varphi_p$ are of type (iii). Since each replacement (iii′) is not the composition of several replacements of type (iii), $p = 1$, and so $\langle \mathcal{B} \rangle$ immediately succeeds $\langle \mathcal{A} \rangle$.

Case 4: (j) *is the replacement* (iv). The statement (2.7) is proved in this case by transposing the matrices in Case 3.

Case 5: (j) *is the replacement* (v)*:*

$$\mathcal{D}_{k_{ij}}(\lambda_i) \oplus \mathcal{D}_{k_{il}}(\lambda_i) \downarrow \mathcal{D}_{k_{ij}-1}(\lambda_i) \oplus \mathcal{D}_{k_{il}+1}(\lambda_i), \quad \text{in which } j < l. \tag{2.11}$$

$\Longrightarrow$. To the contrary, suppose that (2.11) is not (v′); that is, $k_{ij} < k_{i,j+1} < k_{il}$. Then

$$\begin{aligned} \mathcal{D}_{k_{ij}}(\lambda_i) \oplus \mathcal{D}_{k_{i,j+1}}(\lambda_i) \oplus \mathcal{D}_{k_{il}}(\lambda_i) &\downarrow \mathcal{D}_{k_{ij}-1}(\lambda_i) \oplus \mathcal{D}_{k_{i,j+1}+1}(\lambda_i) \oplus \mathcal{D}_{k_{il}}(\lambda_i) \\ &\downarrow \mathcal{D}_{k_{ij}-1}(\lambda_i) \oplus \mathcal{D}_{k_{i,j+1}}(\lambda_i) \oplus \mathcal{D}_{k_{il}+1}(\lambda_i). \end{aligned}$$

By the first main theorem, $\langle\mathcal{B}\rangle$ is not an immediate successor of $\langle\mathcal{A}\rangle$.

$\Longleftarrow$. Let $\mathcal{B}$ be obtained from $\mathcal{A}$ by replacement (v′), and let $\mathcal{B}$ be also obtained from $\mathcal{A}$ by a sequence $\mathcal{A} = \mathcal{A}_1 \overset{\varphi_1}{\longmapsto} \mathcal{A}_2 \overset{\varphi_2}{\longmapsto} \cdots \overset{\varphi_p}{\longmapsto} \mathcal{A}_{p+1} = \mathcal{B}$ of replacements of types (i)–(vi). All replacements $\varphi_1, \dots, \varphi_p$ are not of types (i)–(iv) and (vi) since they change $n_1, n_2, \dots, n_{\overline{s}}$.

Therefore, all $\varphi_1, \dots, \varphi_p$ are of type (v). Since each replacement (v′) is not the composition of several replacements of type (v), $p = 1$, and so $\langle\mathcal{B}\rangle$ immediately succeeds $\langle\mathcal{A}\rangle$.

Case 6: (j) *is the replacement* (vi)*:*

$$\mathcal{L}_{n_i} \oplus \mathcal{L}^T_{m_j} \ \rightsquigarrow\ \mathcal{D}_{r_1}(\mu_1) \oplus \cdots \oplus \mathcal{D}_{r_q}(\mu_q), \tag{2.12}$$

in which $\mu_1, \dots, \mu_q \in \mathbb{C} \cup \infty$ are distinct and $r_1 + \cdots + r_q = n_i + m_j - 1$.

$\Longrightarrow$. To the contrary, suppose that (2.12) is not (vi′).

If $n_i < n_{\overline{s}}$, then

$$\begin{aligned} \mathcal{L}_{n_i} \oplus \mathcal{L}_{n_{\overline{s}}} \oplus \mathcal{L}^T_{m_j} \ &\rightsquigarrow\ \mathcal{L}_{n_i} \oplus \mathcal{D}_{r_1+n_{\overline{s}}-n_i}(\mu_1) \oplus \cdots \oplus \mathcal{D}_{r_q}(\mu_q) \\ &\rightsquigarrow\ \mathcal{L}_{n_{\overline{s}}} \oplus \mathcal{D}_{r_1}(\mu_1) \oplus \cdots \oplus \mathcal{D}_{r_q}(\mu_q), \end{aligned} \tag{2.13}$$

and so $\langle\mathcal{B}\rangle$ is not an immediate successor of $\langle\mathcal{A}\rangle$. Hence $n_i = n_{\overline{s}}$ and, analogously, $m_j = m_{\underline{s}}$.

If there exists $\lambda_i \notin \{\mu_1, \dots, \mu_q\}$, then

$$\begin{aligned} \mathcal{L}_{n_{\overline{s}}} \oplus \mathcal{L}^T_{m_{\underline{s}}} \oplus \mathcal{D}_{k_{i1}}(\lambda_i) \ &\rightsquigarrow\ \mathcal{L}_{n_{\overline{s}}+k_{i1}} \oplus \mathcal{L}^T_{m_{\underline{s}}} \\ &\rightsquigarrow\ \mathcal{D}_{r_1}(\mu_1) \oplus \cdots \oplus \mathcal{D}_{r_q}(\mu_q) \oplus \mathcal{D}_{k_{i1}}(\lambda_i), \end{aligned}$$

and so $\langle\mathcal{B}\rangle$ is not an immediate successor of $\langle\mathcal{A}\rangle$. Hence $q \geqslant t$ and we can rearrange $\mu_1, \dots, \mu_q$ such that $\mu_1 = \lambda_1, \dots, \mu_t = \lambda_t$.

Let $r_i < k_{is_i}$ for some i; for definiteness, for $i = 1$. Then $\mu_1 = \lambda_1$,

$$\begin{aligned} \mathcal{L}_{n_{\overline{s}}} \oplus \mathcal{L}^T_{m_{\underline{s}}} \oplus \mathcal{D}_{k_{1s_1}}(\mu_1) \ &\rightsquigarrow\ \mathcal{L}_{n_{\overline{s}}+k_{1s_1}-r_1} \oplus \mathcal{L}^T_{m_{\underline{s}}} \oplus \mathcal{D}_{r_1}(\mu_1) \\ &\rightsquigarrow\ \mathcal{D}_{r_2}(\mu_2) \oplus \cdots \oplus \mathcal{D}_{r_q}(\mu_q) \oplus \mathcal{D}_{k_{1s_1}}(\mu_1) \oplus \mathcal{D}_{r_1}(\mu_1), \end{aligned}$$

and so $\langle\mathcal{B}\rangle$ is not an immediate successor of $\langle\mathcal{A}\rangle$. Hence, $r_1 \geqslant k_{1s_1}, \ldots, r_t \geqslant k_{ts_t}$.

$\Longleftarrow$. Let $\mathcal{B}$ be obtained from $\mathcal{A}$ by a replacement

$$\varphi: \ \mathcal{L}_{n_{\overline{s}}} \oplus \mathcal{L}^T_{m_{\underline{s}}} \ \downarrow \ \mathcal{D}_{r_1}(\mu_1) \oplus \cdots \oplus \mathcal{D}_{r_q}(\mu_q), \qquad q \geqslant t \tag{2.14}$$

of type (vi′); that is, $\mu_1 = \lambda_1, \ldots, \mu_t = \lambda_t$, and $k_{1s_1} \leqslant r_1, \ldots, k_{ts_t} \leqslant r_t$.

Let $\mathcal{B}$ be also obtained from $\mathcal{A}$ by a sequence $\mathcal{A} = \mathcal{A}_1 \xmapsto{\varphi_1} \mathcal{A}_2 \xmapsto{\varphi_2} \cdots \xmapsto{\varphi_p} \mathcal{A}_{p+1} = \mathcal{B}$ of replacements of types (i)–(vi). Exactly one replacement $\varphi_u : \mathcal{A}_u \to \mathcal{A}_{u+1}$ is of type (vi) since φ decreases $\overline{s}$ by one. The preceding replacements $\varphi_1, \ldots, \varphi_{u-1}$ of types (i)–(v) do not change $\overline{s}$ and $\underline{s}$. Let

$$\mathcal{A}' := \mathcal{A}_u = \bigoplus_{i=1}^{\underline{s}} \mathcal{L}^T_{m'_i} \oplus \bigoplus_{i=1}^{\overline{s}} \mathcal{L}_{n'_i} \oplus \bigoplus_{i=1}^{t'} \Big(\mathcal{D}_{k'_{i1}}(\lambda_i) \oplus \cdots \oplus \mathcal{D}_{k'_{is'_i}}(\lambda_i)\Big),$$
$$m'_1 \leqslant \cdots \leqslant m'_{\underline{s}}, \quad n'_1 \leqslant \cdots \leqslant n'_{\overline{s}}, \quad k'_{i1} \leqslant \cdots \leqslant k'_{is'_i} \ (i = 1, \ldots, t'), \quad t' \leqslant t.$$

We can suppose that φ_u is not a product of replacements. Then φ_u is of type (iv′) due to part "$\Longrightarrow$":

$$\varphi_u: \ \mathcal{L}^T_{m'_{\underline{s}}} \oplus \mathcal{L}_{n'_{\overline{s}}} \ \downarrow \ \mathcal{D}_{\rho_1}(\nu_1) \oplus \cdots \oplus \mathcal{D}_{\rho_{q'}}(\nu_{q'}), \qquad q' \geqslant t', \tag{2.15}$$

in which $\nu_1 = \lambda_1, \ldots, \nu_{t'} = \lambda_{t'}$, and $k_{1s_1} \leqslant \rho_1, \ldots, k_{t's_{t'}} \leqslant \rho_{t'}$.

If $m'_{\underline{s}} > m_{\underline{s}}$, then $m_{\underline{s}}$ has been increased by some φ_l with $l < u$ of type (iv). However, this φ_l decreases $\sum_{i,j} k_{ij}$, which cannot be restored because of the condition $k_{1s_1} \leqslant r_1, \ldots, k_{ts_t} \leqslant r_t$. Hence $m'_{\underline{s}} \leqslant m_{\underline{s}}$ and, analogously, $n'_{\overline{s}} \leqslant n_{\overline{s}}$.

If $m'_{\underline{s}} < m_{\underline{s}}$, then $\sum_{i,j} k'_{ij} + \sum_i \rho_i < \sum_{i,j} k_{ij} + \sum_i r_i$ and this inequality cannot be transformed to the equality by replacements $\varphi_{u+1}, \ldots, \varphi_p$ of types (i)–(v). Hence $m'_{\underline{s}} = m_{\underline{s}}$ and, analogously, $n'_{\overline{s}} = n_{\overline{s}}$.

If $\rho_1 < r_1$, then

$$k'_{11} + \cdots + k'_{1s'_1} + \rho_1 < k_{11} + \cdots + k_{1s_1} + r_1,$$

and this inequality cannot be transformed to the equality by replacements $\varphi_{u+1}, \ldots, \varphi_p$ of types (i)–(v). Hence $\rho_1 \geqslant r_1$ and, analogously, $\rho_i \geqslant r_i$ for all i. Using $m'_{\underline{s}} = m_{\underline{s}}$ and $n'_{\overline{s}} = n_{\overline{s}}$, we find that $t' = t$ and $\rho_i = r_i$ for all i. Therefore, φ_u is the replacement φ from (2.14). It is easy to check that $u = p = 1$ and $\varphi_1 = \varphi$.

2.3. A weak form of Corollary 2.1

Due to the following theorem, which is a weak form of Corollary 2.1, it suffices to describe immediate successors for all pairs (2.6) with two direct summands and for all matrix pairs of the form $\mathcal{D}_{k_1}(\lambda)\oplus\cdots\oplus\mathcal{D}_{k_l}(\lambda)$.

Theorem 2.1. *Let* $\mathcal{A}=([A_{ij}],[A'_{ij}])$ *be the Kronecker pair* (2.6) *partitioned such that the pairs of diagonal blocks* $(A_{11},A'_{11}),(A_{22},A'_{22}),\dots$ *are the direct summands* $\mathcal{L}^T_{m_1},\mathcal{L}^T_{m_2},\dots,\mathcal{L}^T_{m_{\underline{s}}},\mathcal{D}_{k_{11}}(\lambda_1),\mathcal{D}_{k_{12}}(\lambda_1),\dots$ *from* (2.6). *Write*

$$\mathcal{D}_i:=\mathcal{D}_{k_{i1}}(\lambda_1)\oplus\cdots\oplus\mathcal{D}_{k_{is_i}}(\lambda_i),\qquad i=1,\dots,t.$$

Then each immediate successor of $\langle\mathcal{A}\rangle$ *is the orbit of some matrix pair that is obtained from* $\mathcal{A}$ *by an arbitrarily small perturbation of only one matrix pair* (A_{ij},A'_{ij}) *with* $i<j$ *that is not contained in* $\mathcal{D}_1,\dots,\mathcal{D}_t$, *or of only one matrix pair* (A_{ij},A'_{ij}) *from* $\mathcal{D}_1,\dots,\mathcal{D}_t$.

Proof. We consider the partition of the matrices of $\mathcal{A}=(A,A')$ into blocks A_{ij} and A'_{ij} as well as the partition of A and A' into *superblocks*: we join all strips in $([A_{ij}],[A'_{ij}])$ that correspond to the same eigenvalue. Thus, the diagonal superblocks form the pairs

$$\mathcal{L}^T_{m_1},\ \dots,\ \mathcal{L}^T_{m_{\underline{s}}},\ \ \mathcal{D}_1,\dots,\mathcal{D}_t,\ \ \mathcal{L}_{n_1},\ \dots,\ \mathcal{L}_{n_{\overline{s}}}. \tag{2.16}$$

Let $\langle\mathcal{B}\rangle$ be an immediate successor of $\langle\mathcal{A}\rangle$. Then there exists a sequence

$$\mathcal{B}_1=(B_1,B'_1),\ \mathcal{B}_2=(B_2,B'_2),\ \dots \tag{2.17}$$

of pairs from $\langle\mathcal{B}\rangle$ that converges to $\mathcal{A}=(A,A')$. All matrix pairs close enough to $\mathcal{A}$ are reduced to its miniversal normal form (1.19) by a smooth equivalence transformation that preserves $\mathcal{A}$. Hence, all pairs (2.17) can be taken in the miniversal normal form (1.19), which is *upper superblock triangular.*

We say that a block (superblock) of B_i or B'_i in (2.17) is *perturbed* if it differs from the corresponding block (superblock) of A or A'.

Case 1: There are infinite many pairs (2.17), *in which at least one upper diagonal superblock is perturbed.*

Then there is a partition

$$\mathcal{A} = \left(\begin{bmatrix} M & O \\ 0 & N \end{bmatrix}, \begin{bmatrix} M' & O' \\ 0 & N' \end{bmatrix} \right) \quad (O \text{ and } O' \text{ are zero}) \tag{2.18}$$

that is coarser than the partition into superblocks, with the property: O or O' is perturbed infinitely many times in the sequence (2.17). We can suppose that O or O' is perturbed in *each* pair (2.17).

Let $m \times m'$ be the size of (M, M'). Partition

$$\mathcal{B}_i = \left(\begin{bmatrix} M_i & O_i \\ 0 & N_i \end{bmatrix}, \begin{bmatrix} M'_i & O'_i \\ 0 & N'_i \end{bmatrix} \right)$$

conformally with $\mathcal{A}$, write $\xi_i := (\|O_i\| + \|O'_i\|)^{-1}$, and define the equivalent pair

$$\widehat{\mathcal{B}}_i := \begin{bmatrix} I_m & 0 \\ 0 & \xi_i^{-1} I \end{bmatrix} \mathcal{B}_i \begin{bmatrix} I_{m'} & 0 \\ 0 & \xi_i I \end{bmatrix} = \left(\begin{bmatrix} M_i & \xi_i O_i \\ 0 & N_i \end{bmatrix}, \begin{bmatrix} M'_i & \xi_i O'_i \\ 0 & N'_i \end{bmatrix} \right),$$

which belongs to $\langle \mathcal{B} \rangle$. Then $\|\xi_i O_i\| + \|\xi_i O'_i\| = 1$, and so the set of matrix pairs $(\xi_i O_i, \xi_i O'_i)$ is compact. Chose a fundamental subsequence $(\xi_{i_k} O_{i_k}, \xi_{i_k} O'_{i_k})$ and denote its limit by (Q, Q'). Consider the pair

$$\mathcal{X} := \left(\begin{bmatrix} M & Q \\ 0 & N \end{bmatrix}, \begin{bmatrix} M' & Q' \\ 0 & N' \end{bmatrix} \right). \tag{2.19}$$

We have $\langle \mathcal{B} \rangle \geqslant \langle \mathcal{X} \rangle$ since $\widehat{\mathcal{B}}_{i_k} \to \mathcal{X}$ as $k \to \infty$ and all $\widehat{\mathcal{B}}_{i_k} \in \langle \mathcal{B} \rangle$.

Make additional partitions of $\mathcal{X}$ into blocks conformally to the partition of $\mathcal{A} = ([A_{ij}], [A'_{ij}])$ in the theorem. Choose in (Q, Q') the nonzero pair (X, X') of conformal blocks X and X' such that all columns of Q to the left of X and all blocks of Q exactly under X are zero, and all columns of

Q' to the left of X' and all blocks of Q' exactly under X' are zero:

$$\mathcal{X} = \left(\left[\begin{array}{ccc|ccc} M_1 & & 0 & 0 & * & * \\ & M_2 & & 0 & X & * \\ 0 & & M_3 & 0 & 0 & * \\ \hline & & & N_1 & & 0 \\ & 0 & & & N_2 & \\ & & & 0 & & N_3 \end{array}\right], \left[\begin{array}{ccc|ccc} M_1' & & 0 & 0 & * & * \\ & M_2' & & 0 & X' & * \\ 0 & & M_3' & 0 & 0 & * \\ \hline & & & N_1' & & 0 \\ & 0 & & & N_2' & \\ & & & 0 & & N_3' \end{array}\right] \right).$$

Write

$$\mathcal{Y} = \left(\left[\begin{array}{c|c} M & Y \\ \hline 0 & N \end{array}\right], \left[\begin{array}{c|c} M' & Y' \\ \hline 0 & N' \end{array}\right] \right)$$

$$:= \left(\left[\begin{array}{ccc|ccc} M_1 & & 0 & 0 & 0 & 0 \\ & M_2 & & 0 & X & 0 \\ 0 & & M_3 & 0 & 0 & 0 \\ \hline & & & N_1 & & 0 \\ & 0 & & & N_2 & \\ & & & 0 & & N_3 \end{array}\right], \left[\begin{array}{ccc|ccc} M_1' & & 0 & 0 & 0 & 0 \\ & M_2' & & 0 & X' & 0 \\ 0 & & M_3' & 0 & 0 & 0 \\ \hline & & & N_1' & & 0 \\ & 0 & & & N_2' & \\ & & & 0 & & N_3' \end{array}\right] \right). \tag{2.20}$$

Then

$$(I_a \oplus \varepsilon^{-1} I \oplus \varepsilon^{-2} I_c)\mathcal{X}(I_b \oplus \varepsilon I \oplus \varepsilon^2 I_d) \xrightarrow{\text{as } \varepsilon\to 0} \mathcal{Y},$$

in which $a \times b$ is the size of (M_1, M_1') and $c \times d$ is the size of (N_3, N_3'). This implies that $\langle \mathcal{X} \rangle \geqslant \langle \mathcal{Y} \rangle$. Since

$$\mathcal{Y}_\varepsilon := \begin{bmatrix} I_m & 0 \\ 0 & \varepsilon^{-1} I \end{bmatrix} \mathcal{Y} \begin{bmatrix} I_{m'} & 0 \\ 0 & \varepsilon I \end{bmatrix} = \left(\begin{bmatrix} M & \varepsilon Y \\ 0 & N \end{bmatrix}, \begin{bmatrix} M' & \varepsilon Y' \\ 0 & N' \end{bmatrix} \right) \xrightarrow{\text{as } \varepsilon\to 0} \mathcal{A},$$

we have that $\langle \mathcal{Y} \rangle \geqslant \langle \mathcal{A} \rangle$. Therefore, $\langle \mathcal{B} \rangle \geqslant \langle \mathcal{X} \rangle \geqslant \langle \mathcal{Y} \rangle \geqslant \langle \mathcal{A} \rangle$.

In order to prove that $\langle \mathcal{Y} \rangle$ is a desired pair, it suffices to prove that $\langle \mathcal{Y} \rangle \neq \langle \mathcal{A} \rangle$. Indeed, then $\langle \mathcal{B} \rangle = \langle \mathcal{Y} \rangle > \langle \mathcal{A} \rangle$ because $\langle \mathcal{B} \rangle$ is an immediate successor of $\langle \mathcal{A} \rangle$.

On the contrary, suppose that $\langle \mathcal{Y} \rangle = \langle \mathcal{A} \rangle$. Then $\mathcal{Y}_\varepsilon \in \langle \mathcal{A} \rangle$ for each ε since $\mathcal{Y}_\varepsilon \sim \mathcal{Y}$. Hence there exist nonsingular matrices, which we take in the form

$I + R_\varepsilon$ and $I + S_\varepsilon$, such that

$$\mathcal{Y}_\varepsilon = (I + R_\varepsilon)\mathcal{A}(I + S_\varepsilon) = \mathcal{A} + R_\varepsilon\mathcal{A} + \mathcal{A}S_\varepsilon + R_\varepsilon\mathcal{A}S_\varepsilon. \tag{2.21}$$

By Lipschitz's property for matrix pairs (see [61] or [1]), we can chose the matrices R_ε, S_ε and a positive constant $c \in \mathbb{R}$ such that

$$\|R_\varepsilon\| < \varepsilon c, \qquad \|S_\varepsilon\| < \varepsilon c \tag{2.22}$$

for all ε, in which $\|\cdot\|$ is the Frobenius matrix norm.

The pair $\mathcal{Y}_\varepsilon$ is in miniversal normal form of (2.18) since all nonzero entries of Q and Q' are at the places of some stars in (1.19). By the construction of the miniversal deformation in [39],

$$\Delta\mathcal{Y}_\varepsilon := \mathcal{Y}_\varepsilon - \mathcal{A} = \varepsilon\left(\left[\begin{array}{c|c} 0 & \begin{smallmatrix} 0 & 0 & 0 \\ 0 & X & 0 \\ 0 & 0 & 0 \end{smallmatrix} \\ \hline 0 & 0 \end{array}\right], \left[\begin{array}{c|c} 0 & \begin{smallmatrix} 0 & 0 & 0 \\ 0 & X' & 0 \\ 0 & 0 & 0 \end{smallmatrix} \\ \hline 0 & 0 \end{array}\right]\right) = R_\varepsilon\mathcal{A} + \mathcal{A}S_\varepsilon + R_\varepsilon\mathcal{A}S_\varepsilon \tag{2.23}$$

does not belong to the space

$$\mathbb{T} := \{R\mathcal{A} + \mathcal{A}S \,|\, R \text{ and } S \text{ are nonsingular matrices}\}.$$

Thus,

$$d_\varepsilon := \min\left\{\|\mathcal{Y}_\varepsilon - \mathcal{A} - R\mathcal{A} - \mathcal{A}S\| \,\middle|\, R, S \text{ are square matrices}\right\} \neq 0. \tag{2.24}$$

(Note that $\mathbb{T}$ is the tangent space at $\mathcal{A}$ to the orbit of $\mathcal{A}$, and d_ε is the distance from $\mathcal{Y}_\varepsilon$ to the affine space $\{\mathcal{A} + R\mathcal{A} + \mathcal{A}S \,|\, R, S\}$.)

Let R' and S' be such that

$$d_1 = \|\mathcal{Y}_1 - \mathcal{A} - R'\mathcal{A} - \mathcal{A}S'\| = \|\Delta\mathcal{Y}_1 - R'\mathcal{A} - \mathcal{A}S'\|.$$

By (2.23), $\Delta\mathcal{Y}_\varepsilon = \varepsilon\Delta\mathcal{Y}_1$, and so $\varepsilon d_1 = \|\Delta\mathcal{Y}_\varepsilon - (\varepsilon R')\mathcal{A} - \mathcal{A}(\varepsilon S')\| = d_\varepsilon$. By (2.22),

$$\varepsilon d_1 \leqslant \|\Delta\mathcal{Y}_\varepsilon - R_\varepsilon\mathcal{A} - \mathcal{A}S_\varepsilon\| = \|R_\varepsilon\mathcal{A}S_\varepsilon\| \leqslant \|R_\varepsilon\|\|\mathcal{A}\|\|S_\varepsilon\| \leqslant \varepsilon^2c^2\|\mathcal{A}\|.$$

This leads to a contradiction since $\varepsilon d_1 \leqslant \varepsilon^2c^2\|\mathcal{A}\|$ does not hold for a sufficiently small ε.

Case 2: There is only a finite number of pairs (2.17) *in which at least one upper diagonal superblock is perturbed.*

Write $\mathcal{A} = \mathcal{A}^{(1)} \oplus \mathcal{A}^{(2)} \oplus \cdots$, in which $\mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \dots$ are the pairs of diagonal superblocks of $\mathcal{A}$. We can suppose that all upper diagonal superblocks are not perturbed. Then $\mathcal{B}_i := \mathcal{B}_i^{(1)} \oplus \mathcal{B}_i^{(2)} \oplus \cdots$, in which $\mathcal{B}_i^{(1)}, \mathcal{B}_i^{(2)}, \dots$ are the pairs of perturbed diagonal superblocks of $\mathcal{B}_i$ in (2.17).

Since all $\mathcal{B}_i \sim \mathcal{B}$, we can suppose that $\mathcal{B}_1^{(l)} \sim \mathcal{B}_2^{(l)} \sim \cdots$ for each l. Since $\mathcal{A} \nsim \mathcal{B}$, $\mathcal{A}^{(l)} \nsim \mathcal{B}_1^{(l)} \sim \mathcal{B}_2^{(l)} \sim \cdots$ for some l. Then all

$$\mathcal{C}_i := \mathcal{A}^{(1)} \oplus \cdots \oplus \mathcal{A}^{(l-1)} \oplus \mathcal{B}_i^{(l)} \oplus \mathcal{A}^{(l+1)} \oplus \cdots$$

are equivalent and their orbit $\langle \mathcal{C}_1 \rangle > \langle \mathcal{A} \rangle$. Moreover, $\langle \mathcal{B} \rangle \geqslant \langle \mathcal{C}_1 \rangle$ because

$$\mathcal{B}_i^{(1)} \oplus \cdots \oplus \mathcal{B}_i^{(l-1)} \oplus \mathcal{B}_1^{(l)} \oplus \mathcal{B}_i^{(l+1)} \oplus \cdots \xrightarrow{\text{as } i \to \infty} \mathcal{C}_1.$$

Since there is no intermediate orbit between $\langle \mathcal{A} \rangle$ and $\langle \mathcal{B} \rangle$, we have that $\langle \mathcal{B} \rangle = \langle \mathcal{C}_1 \rangle$. $\square$

2.4. Perturbations of direct sums of two indecomposable Kronecker pairs

2.4.1. Perturbations of $\mathcal{L}_m^T \oplus \mathcal{L}_n^T$

Theorem 2.2. (a) *The set of Kronecker's canonical forms of all pairs in a sufficiently small neighborhood of*

$$\mathcal{L}_m^T \oplus \mathcal{L}_n^T, \qquad m \leqslant n \tag{2.25}$$

consists of the pairs

$$\mathcal{L}_{m+r}^T \oplus \mathcal{L}_{n-r}^T, \qquad m + r \leqslant n - r,\ r \geqslant 0. \tag{2.26}$$

(b) *Each pair* (2.26) *with $r > 0$ is equivalent to a pair of the form*

$$\left(\begin{bmatrix} L_m^T & 0 \\ 0 & L_n^T \end{bmatrix}, \begin{bmatrix} R_m^T & \Delta_r(\varepsilon) \\ 0 & R_n^T \end{bmatrix}\right) \tag{2.27}$$

(which is obtained by an arbitrarily small perturbation of (2.25)*), in which $\Delta_r(\varepsilon)$ is defined in* (2.5) *and ε is an arbitrary nonzero complex number.*

Lemma 2.1. *Each pair of $n \times (n-1)$ matrices of the form*

$$\left(\begin{bmatrix} 1 & * & & * \\ 0 & 1 & \ddots & \\ & 0 & \ddots & * \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix}, \begin{bmatrix} * & * & & * \\ 1 & * & \ddots & \\ & 1 & \ddots & * \\ & & \ddots & * \\ 0 & & & 1 \end{bmatrix}\right) \tag{2.28}$$

is reduced to $\mathcal{L}_n^T$ by simultaneous additions of columns from left to right and simultaneous additions of rows from the bottom to up.

Proof. Consider the subpair $\mathcal{P}$ of (2.28) obtained by removing the last row and last column in the matrices of the pair (2.28). We reduce (2.28) by simultaneous additions of columns of its matrices from left to right and simultaneous additions of rows from the bottom to up. Reasoning by induction on n, we reduce the subpair $\mathcal{P}$ to $\mathcal{L}_{n-1}^T$ and obtain the new (2.28) in which all stars are zero except for some stars of the last columns. We make zero the stars of the last column in the first matrix by adding the other columns simultaneously in both matrices; then we make zero the stars of the last column in the second matrix by adding the last row. $\square$

Proof of Theorem 3.48. (a) By Theorem 1.3, there is a neighborhood of (2.25), in which all pairs are equivalent to pairs of the form

$$(C,D) := \left(\left[\begin{array}{c|c} L_m^T & 0 \\ \hline 0 & L_n^T \end{array}\right], \left[\begin{array}{c|c} R_m^T & \begin{matrix} \alpha_1 \ \dots \ \alpha_{n-1} \\ 0 \end{matrix} \\ \hline 0 & R_n^T \end{array}\right] \right), \quad \text{all } \alpha_i \in \mathbb{C}, \tag{2.29}$$

in which the last m entries in the sequence $\alpha_1, \dots, \alpha_{n-1}$ are zero. It is sufficient to prove that (C,D) is equivalent to a pair of the form (2.26).

If $\alpha_1 = \dots = \alpha_{n-1} = 0$, then (C,D) is the pair (2.25). Let α_s be the first nonzero entry, and so

$$1 \leqslant s < n-m \quad \text{if } m \neq n. \tag{2.30}$$

Let us reduce (C,D) by simultaneous elementary transformations to the form (2.26). We usually specify only transformations with one of the matrices C and D which means that we make the same transformations with the other matrix. We divide the first horizontal strips of C and D by α_s, then multiply the first vertical strips by α_s, and obtain

$$
(C,D)=\left(\left[\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22}\end{array}\right],\left[\begin{array}{c|c} D_{11} & D_{12} \\ \hline D_{21} & D_{22}\end{array}\right]\right)
$$

$$
=\left(\begin{array}{c}
\begin{array}{c|ccc} \scriptstyle 1\ \ m-1 & \scriptstyle 1 \quad s & & \scriptstyle s+m-1 \end{array}\\
\left[\begin{array}{c|c:c:c}
\begin{smallmatrix}1&&\\0&\ddots&\\&\ddots&1\\&&0\end{smallmatrix} & & & \\ \hline
 & \begin{smallmatrix}1&&&\\0&\ddots&&\\&\ddots&1&\\&&0&1\end{smallmatrix} & & \\ \hdashline
\begin{smallmatrix}\not{0}&&\\0&\ddots&\\&\ddots&\not{0}\\&&0\end{smallmatrix} & \begin{smallmatrix}&0\\&\end{smallmatrix} & \boxed{\begin{smallmatrix}1&&\\0&\ddots&\\&\ddots&1\\&&0\end{smallmatrix}} & \begin{smallmatrix}\\1\end{smallmatrix} \\ \hdashline
 & & & \begin{smallmatrix}0&\ddots&\\&\ddots&1\\&&0\end{smallmatrix}
\end{array}\right]
\end{array},\;
\begin{array}{c}
\begin{array}{c|ccc} \scriptstyle 1\ \ m-1 & \scriptstyle 1 \quad s & & \scriptstyle s+m-1 \end{array}\\
\left[\begin{array}{c|c:c:c}
\begin{smallmatrix}0&&\\1&\ddots&\\&\ddots&0\\&&1\end{smallmatrix} & 0\dots 0\,1 & *\ \dots\ * & *\dots* \\ \hline
 & \begin{smallmatrix}0&&&\\1&\ddots&&\\&\ddots&0&\\&&1&0\end{smallmatrix} & & \\ \hdashline
\begin{smallmatrix}0&&\\ \not{0}&\ddots&\\&\ddots&0\\&&\not{0}\end{smallmatrix} & \begin{smallmatrix}&1\\&\end{smallmatrix} & \boxed{\begin{smallmatrix}0&&\\1&\ddots&\\&\ddots&0\\&&1\end{smallmatrix}} & \begin{smallmatrix}\\0\end{smallmatrix} \\ \hdashline
 & & & \begin{smallmatrix}1&\ddots&\\&\ddots&0\\&&1\end{smallmatrix}
\end{array}\right]
\end{array}
\begin{array}{l} \scriptstyle 1\\ \scriptstyle m\\ \scriptstyle 1\\ \scriptstyle s\\ \scriptstyle s+1\\ \scriptstyle s+m \end{array}
\right) \tag{2.31}
$$

with $\alpha_s=1$. We reduce (C,D) by the following simultaneous elementary transformations in order to make zero the entry "1" under α_s (the zero entries in (2.31) that are transformed to -1 and then are restored to 0 are denoted by $\not{0}$):

- The strip $[D_{11}\ D_{12}]$ is subtracted from the substrip formed by rows $s+1,s+2,\dots,s+m$ in the strip $[D_{21}\ D_{22}]$. Thus, the block $(1,1)$ is subtracted from the rectangle (see (2.31)) in the block $(2,1)$.
- Then the substrip formed by columns $s+1,\dots,s+m-1$ in $\left[\begin{smallmatrix}D_{12}\\D_{22}\end{smallmatrix}\right]$ is added to $\left[\begin{smallmatrix}D_{11}\\D_{21}\end{smallmatrix}\right]$. Thus, the rectangle in the block $(2,2)$ is added to the rectangle in the block $(2,1)$ restoring it.

We obtain

$$
(C,D)=\left(\left[\begin{array}{c|c|c}
\begin{smallmatrix}1&&\\0&\ddots&\\&\ddots&1\\&&0\end{smallmatrix} & & \\ \hline
 & \begin{smallmatrix}1&&&\\&\ddots&&\\&&1&\\&&&1\end{smallmatrix} & \\ \hline
 & & \begin{smallmatrix}1&&\\0&\ddots&\\&\ddots&1\\&&0\end{smallmatrix}
\end{array}\right],
\left[\begin{array}{c|c|c}
\begin{smallmatrix}*&\dots&*\\1&&0\\&\ddots&\\&&1\end{smallmatrix} & 0\ \dots\ 0\ \ 1 & *\ \dots\ * \\ \hline
 & \begin{smallmatrix}0&&&\\1&\ddots&&\\&\ddots&0&\\&&1&0\end{smallmatrix} & \\ \hline
*\ \dots\ * & & \begin{smallmatrix}*&\dots&*\\1&&0\\&\ddots&\\&&1\end{smallmatrix}
\end{array}\right]\right), \tag{2.32}
$$

in which the stars denote complex numbers. Interchange the first and second vertical strips, then the first and second horizontal strips, and obtain

$$
(C,D)=\left(\left[\begin{array}{c|c|c} C_{11} & C_{12} & C_{13} \\ \hline C_{21} & C_{22} & C_{23} \\ \hline C_{31} & C_{32} & C_{33} \end{array}\right],\ \left[\begin{array}{c|c|c} D_{11} & D_{12} & D_{13} \\ \hline D_{21} & D_{22} & D_{23} \\ \hline D_{31} & D_{32} & D_{33} \end{array}\right]\right)
$$

$$
=\left(\left[\begin{array}{cccc|cccc|ccc}
1 &&&&&&&&&& \\
0 & \ddots &&&&&&&&& \\
& \ddots & 1 &&&&&&&& \\
&& 0 & 1 &&&&&&& \\ \hline
&&& 0 & 1 &&&&&& \\
&&&& 0 & 1 &&&&& \\
&&&&& 0 & \ddots &&&& \\
&&&&&& \ddots & 1 &&& \\
&&&&&&& 0 &&& \\ \hline
&&&& 0 & * & \dots & * & 1 && \\
&&&&& 0 & \ddots & \vdots & 0 & \ddots & \\
&&&&&& \ddots & * && \ddots & 1 \\
&&&&&&& 0 &&& 0
\end{array}\right],\ \left[\begin{array}{cccc|cccc|ccc}
0 &&&&&&&&&& \\
1 & \ddots &&&&&&&&& \\
& \ddots & 0 &&&&&&&& \\
&& 1 & 0 &&&&&&& \\ \hline
&&& 1 & * & * & \dots & * & * & \dots & * \\
&&&& 1 &&& 0 &&& \\
&&&&& 1 &&&&& \\
&&&&&& \ddots &&&& \\
&&&&&&& 1 &&& \\ \hline
&&&& * & * & \dots & * & * & \dots & * \\
&&&&& * & \ddots & \vdots & 1 && 0 \\
&&&&&& \ddots & * && \ddots & \\
&&&&&&& * &&& 1
\end{array}\right]\right), \qquad (2.33)
$$

in which we replace by stars some zero entries of blocks C_{32} and D_{32}.

Using transformations from Lemma 2.1, we make zero all stars in D_{33}; the forms of the other blocks do not change. Make zero row 1 of D_{32} by adding rows $2, 3, \dots$ of horizontal strip 2 to row 1 of strip 3 simultaneously in C and D. Make zero row 1 of C_{32} by adding column 1 of vertical strip 3 simultaneously in C and D. Then, adding rows $3, 4, \dots$ of strip 2 to the row 2 of strip 3, we make zero row 2 of D_{32}. Adding column 2 of vertical strip 3 we make zero row 2 of C_{32}, and so on until we obtain (2.33) in which all stars in horizontal strips 3 of C and D are.

Using Lemma 2.1, we make zero all stars in D_{22}. Multiplying horizontal strips 2 in C and D by an arbitrarily small number and then dividing vertical strips 2 by the same number, we make the entries of D_{23} arbitrarily small; these transformations do not change the other blocks. We obtain the pair that is equivalent to the initial perturbed pair (2.29) and that is obtained from $\mathcal{L}^T_{m+s} \oplus \mathcal{L}^T_{n-s}$ by an arbitrarily small perturbation, in which s is from (2.31) and satisfies (2.30). We interchange $\mathcal{L}^T_{m+s}$ and $\mathcal{L}^T_{n-s}$ if $m+s > n-s$, and reduce the obtained pair by equivalence transformations to its miniversal form

$$\left(\left[\begin{array}{c|c} L^T_{m'} & 0 \\ \hline 0 & L^T_{n'} \end{array}\right], \left[\begin{array}{c|c} R^T_{m'} & \begin{matrix} * \cdots * \\ 0 \end{matrix} \\ \hline 0 & R^T_{n'} \end{array}\right]\right), \tag{2.34}$$

in which the stars are sufficiently small complex numbers. By (2.30),

$$m < m' := \min(m+s, n-s) \leqslant n' := \max(m+s, n-s).$$

We repeat this procedure until we obtain a pair

$$\left(\left[\begin{array}{c|c} L^T_{m^{(l)}} & 0 \\ \hline 0 & L^T_{n^{(l)}} \end{array}\right], \left[\begin{array}{c|c} R^T_{m^{(l)}} & \begin{matrix} * \cdots * \\ 0 \end{matrix} \\ \hline 0 & R^T_{n^{(l)}} \end{array}\right]\right) \tag{2.35}$$

in which all stars are zero, and $m < m^{(l)} \leqslant n^{(l)}$. Thus, (2.35) is of the form (2.26) with $r > 0$.

(b) Let $\mathcal{L}^T_{m+r} \oplus \mathcal{L}^T_{n-r}$ be the pair (2.26) with $r > 0$; we must prove that it is equivalent to (2.27). We divide the first horizontal strips of (2.27) by ε, then multiply the first vertical strips by ε, and obtain the pair (2.31) in which all stairs are zero. The obtained pair is reduced as above to (3.1.1) in which all stairs are zero. This pair is permutation equivalent to $\mathcal{L}^T_{m+r} \oplus \mathcal{L}^T_{n-r}$. □

2.4.2. Perturbations of $\mathcal{L}_n \oplus \mathcal{L}_m$

Theorem 2.3. (a) *The set of Kronecker's canonical forms of all pairs in a sufficiently small neighborhood of*

$$\mathcal{L}_m \oplus \mathcal{L}_n, \qquad m \leqslant n \tag{2.36}$$

consists of the pairs

$$\mathcal{L}_{m+r} \oplus \mathcal{L}_{n-r}, \qquad m + r \leqslant n - r,\ r \geqslant 0. \tag{2.37}$$

(b) *Each pair* (2.37) *with* $r > 0$ *is equivalent to a pair of the form*

$$\left(\begin{bmatrix} L_m & 0 \\ 0 & L_n \end{bmatrix}, \begin{bmatrix} R_m & 0 \\ \Delta_r(\varepsilon)^T & R_n \end{bmatrix}\right)$$

(which is obtained by an arbitrarily small perturbation of (2.36)*), in which* $\Delta_r(\varepsilon)$ *is defined in* (2.5) *and* ε *is an arbitrary nonzero complex number.*

Proof. This theorem is obtained from Theorem 3.48 by matrix transposition. □

2.4.3. Perturbations of $\mathcal{L}_m^T \oplus \mathcal{D}_n(\lambda)$

Theorem 2.4. *The set of Kronecker's canonical forms of all pairs obtained by perturbations of the blocks* $(1,2)$ *in*

$$\mathcal{L}_m^T \oplus \mathcal{D}_n(\lambda) = \begin{cases} \left(\begin{bmatrix} L_m^T & 0 \\ 0 & I_n \end{bmatrix}, \begin{bmatrix} R_m^T & 0 \\ 0 & J_n(\lambda) \end{bmatrix}\right) & \text{if } \lambda \in \mathbb{C} \\ \left(\begin{bmatrix} L_m^T & 0 \\ 0 & J_n(0) \end{bmatrix}, \begin{bmatrix} R_m^T & 0 \\ 0 & I_n \end{bmatrix}\right) & \text{if } \lambda = \infty \end{cases} \tag{2.38}$$

consists of the pairs

$$\mathcal{L}_{m+r}^T \oplus \mathcal{D}_{n-r}(\lambda), \quad \text{in which } 0 \leqslant r \leqslant n. \tag{2.39}$$

(b) *Each pair* (2.39) *with* $r > 0$ *is equivalent to a pair of the form*

$$\begin{matrix} \left(\begin{bmatrix} L_m^T & 0 \\ 0 & I_n \end{bmatrix}, \begin{bmatrix} R_m^T & \Delta_{n-r+1}(\varepsilon) \\ 0 & J_n(\lambda) \end{bmatrix}\right) & \text{if } \lambda \in \mathbb{C} \\ \left(\begin{bmatrix} L_m^T & \nabla_{n-r+1}(\varepsilon) \\ 0 & J_n(0) \end{bmatrix}, \begin{bmatrix} R_m^T & 0 \\ 0 & I_n \end{bmatrix}\right) & \text{if } \lambda = \infty \end{matrix} \tag{2.40}$$

(which is obtained by an arbitrarily small perturbation of (2.38)*), in which* $\Delta_r(\varepsilon)$ *and* $\nabla_r(\varepsilon)$ *are defined in* (2.5) *and* ε *is an arbitrary nonzero complex number.*

Proof. Let (A,B) be the pair (2.38) with $\lambda=\infty$. Since

$$(R_m^T,L_m^T)=Z_m(L_m^T,R_m^T)Z_{m-1},\qquad Z_p:=\begin{bmatrix}0&&1\\&\iddots&\\1&&0\end{bmatrix}\ (p\text{-by-}p),$$

(B,A) is equivalent to the pair (2.38) with $\lambda=0$. Therefore, it suffices to prove the theorem for $\lambda\in\mathbb{C}$.

Let $(A,B(\lambda))$ be the pair (2.38) with $\lambda\in\mathbb{C}$. Since $(L_m^T,R_m^T-\lambda L_m^T)$ is equivalent to (L_m^T,R_m^T), the pair $(A,B(\lambda)-\lambda A)$ is equivalent to $(A,B(0))$. Therefore, it suffices to prove the theorem for $\lambda=0$. In the rest of the proof, we set $\lambda=0$.

(a) Let (C,D) be a pair that is obtained from (2.38) with $\lambda=0$ by replacing its blocks $(1,2)$ by arbitrary matrices; we must prove that the Kronecker canonical form of (C,D) is (2.39) for some r.

Multiplying the first horizontal strips of C and D by an arbitrarily small number and then dividing the first vertical strips by the same number, we make the entries of the blocks $(1,2)$ arbitrarily small. By Theorem 1.3, (C,D) is reduced by equivalence transformations to the form

$$(C,D)=\left(\left[\begin{array}{c|c}C_{11}&C_{12}\\\hline C_{21}&C_{22}\end{array}\right],\left[\begin{array}{c|c}D_{11}&D_{12}\\\hline D_{21}&D_{22}\end{array}\right]\right)$$
$$=\left(\left[\begin{array}{cccc|cccc}1&&&&&&&\\0&\ddots&&&&&&\\&\ddots&1&&&0&&\\&&0&&&&&\\\hline &&&&1&&&\\&&&&&1&&\\&0&&&&&\ddots&\\&&&&&&&1\end{array}\right],\left[\begin{array}{cccc|cccc}0&&&&\alpha_1&\alpha_2&\dots&\alpha_n\\1&\ddots&&&&&&\\&\ddots&0&&&0&&\\&&1&&&&&\\\hline &&&&0&1&&\\&&&&&0&\ddots&\\&0&&&&&\ddots&1\\&&&&&&&0\end{array}\right]\right),\qquad(2.41)$$

in which $\alpha_1,\dots,\alpha_n$ are arbitrarily small.

Each matrix that commutes with $J_n(0)$ has the form

$$K := \begin{bmatrix} \kappa_1 & \kappa_2 & \ddots & \kappa_n \\ & \kappa_1 & \ddots & \ddots \\ & & \ddots & \kappa_2 \\ 0 & & & \kappa_1 \end{bmatrix}, \qquad \kappa_1, \dots, \kappa_n \in \mathbb{C}.$$

The equivalence transformation

$$(I_m \oplus K^{-1})(C, D)(I_{m-1} \oplus K), \qquad c_1 \neq 0$$

replaces $(\alpha_1, \dots, \alpha_n)$ by

$$(\alpha_1, \dots, \alpha_n)K = (\alpha_1\kappa_1,\ \alpha_1\kappa_2 + \alpha_2\kappa_1,\ \dots,\ \alpha_1\kappa_n + \dots + \alpha_n\kappa_1) \tag{2.42}$$

and does not change the other entries of C and D. Let α_s be the first nonzero entry in $(\alpha_1, \dots, \alpha_n)$. Using transformations (2.42), we make $(\alpha_1, \dots, \alpha_n) = (0, \dots, 0, 1, 0, \dots, 0)$ with "1" at the position s.

Case 1: $s = 1$. The pair (C, D) is permutation equivalent to (L_{m+n}^T, R_{m+n}^T), which is a pair of the form (2.39).

Case 2: $s \geqslant 2$. The "1" under $\alpha_s = 1$ is the $(s-1, s)$th entry of the block D_{22} (see (2.41)). We make zero this entry of D_{22} by the following elementary transformations:

- Let first $m < s$. We subtract the rows $1, 2, \dots, m$ of the first horizontal strip from the rows $s-1, s-2, \dots, s-m$ of the second horizontal strip, respectively, in C and D. Then we add the columns $s-1, s-2, \dots, s-m+1$ of the second vertical strip to the columns $1, 2, \dots, m-1$ of the first vertical strip in C and D. For example, if $m = 3$, $n = 6$, and $s = 5$, then

$$(C, D) = \left(\left[\begin{array}{cc|cc:cc:cc} 1 & 0 & & & & & & \\ 0 & 1 & & & & & & \\ 0 & 0 & & & & & & \\ \hline & & 1 & & & & & \\ \hdashline 0 & 0 & & 1 & 0 & 0 & & \\ 0 & \cancel{0} & & & 1 & 0 & & \\ \cancel{0} & 0 & & & 0 & 1 & & \\ \hdashline & & & & & & 1 & \\ & & & & & & & 1 \end{array}\right], \left[\begin{array}{cc|cc:cc:cc} 0 & 0 & & & & & 1 & \\ 1 & 0 & & & & & & \\ 0 & 1 & & & & & & \\ \hline & & 0 & 1 & & & & \\ \hdashline 0 & \cancel{0} & & 0 & 1 & 0 & & \\ \cancel{0} & 0 & & & 0 & 1 & & \\ 0 & 0 & & & 0 & 0 & 1 & \\ \hdashline & & & & & & 0 & 1 \\ & & & & & & & 0 \end{array}\right] \right);$$

the zero entries that are transformed to -1 and then are restored to 0 are denoted by $\cancel{0}$.

- Let now $m \geqslant s$. We subtract the rows $1, 2, \dots, s-1$ of the first horizontal strip from the rows $s-1, s-2, \dots, 1$ of the second horizontal strip, respectively, in C and D. Then we add the columns $s-1, s-2, \dots, 1$ of the second vertical strip to the columns $1, 2, \dots, s-1$ of the first vertical strip in C and D. For example, if $m = 5$, $n = 4$, and $s = 3$, then

$$(C, D) = \left(\left[\begin{array}{cc:cc|cc:cc}
1 & 0 & & & & & & \\
0 & 1 & & & & & & \\ \hline
 & 0 & 1 & & & & & \\
 & & 0 & 1 & & & & \\
 & & & 0 & & & & \\ \hline
0 & \cancel{0} & & & 1 & 0 & & \\
\cancel{0} & 0 & & & 0 & 1 & & \\ \hdashline
 & & & & & & 1 & \\
 & & & & & & & 1
\end{array}\right], \left[\begin{array}{cc:cc|cc:cc}
0 & 0 & & & & & 1 & \\
1 & 0 & & & & & & \\ \hline
 & 1 & 0 & & & & & \\
 & & 1 & 0 & & & & \\
 & & & 1 & & & & \\ \hline
\cancel{0} & 0 & & & 0 & 1 & & \\
0 & 0 & & & 0 & 0 & 1 & \\ \hdashline
 & & & & & & 0 & 1 \\
 & & & & & & & 0
\end{array}\right] \right).$$

In both the cases, we have reduced (C, D) to the pair that is obtained from (2.41) by replacing $(\alpha_1, \dots, \alpha_n)$ by $(0, \dots, 0, 1, 0, \dots, 0)$ and the entry "1" under α_s by 0. This pair is permutation equivalent to $(L^T_{m+n-s+1}, R^T_{m+n-s+1}) \oplus (I_{s-1}, J_{s-1}(0))$, which is a pair of the form (2.39).

(b) Let $(E, F) := \mathcal{L}^T_{m+r} \oplus \mathcal{D}_{n-r}(0)$ with $0 < r \leqslant n$ be the pair (2.39) with $\lambda = 0$; we must prove that it is equivalent to (2.40). The pair (2.40) with $\lambda = 0$ is the pair (2.41) in which $(\alpha_1, \dots, \alpha_n) = (0, \dots, 0, \varepsilon, 0, \dots, 0)$ and $\varepsilon \neq 0$ at the place $s := n - r + 1$. Reasoning as in Cases 1 and 2 of the part (a), we reduce it to a pair that is permutation equivalent to (E, F). $\square$

2.4.4. Perturbations of $\mathcal{L}_m \oplus \mathcal{D}_n(\lambda)$

Theorem 2.5. (a) *The set of Kronecker's canonical forms of all pairs obtained by perturbations of the blocks* $(2, 1)$ *in*

$$\mathcal{L}_m \oplus \mathcal{D}_n(\lambda) \tag{2.43}$$

consists of the pairs

$$\mathcal{L}_{m+r} \oplus \mathcal{D}_{n-r}(\lambda), \quad \textit{in which } 0 \leqslant r \leqslant n. \tag{2.44}$$

(b) *Each pair* (2.44) *with* $r > 0$ *is equivalent to a pair of the form*

$$\begin{aligned} &\left(\begin{bmatrix} L_m & 0 \\ 0 & I_n \end{bmatrix}, \begin{bmatrix} R_m & 0 \\ \Delta_r(\varepsilon)^T & J_n(\lambda) \end{bmatrix} \right) \quad \textit{if } \lambda \in \mathbb{C} \\ &\left(\begin{bmatrix} L_m & 0 \\ 0 & J_n(0) \end{bmatrix}, \begin{bmatrix} R_m & 0 \\ \nabla_r(\varepsilon)^T & I_n \end{bmatrix} \right) \quad \textit{if } \lambda = \infty \end{aligned}$$

(which is obtained by an arbitrarily small perturbation of (2.43)*), in which* ε *is an arbitrary nonzero complex number.*

Proof. The mapping

$$A \mapsto \begin{bmatrix} I_{m-1} & 0 \\ 0 & Z_n \end{bmatrix} A^T \begin{bmatrix} I_m & 0 \\ 0 & Z_n \end{bmatrix}, \qquad Z_n := \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix} \ (n\text{-by-}n)$$

transforms the matrices from Theorem 2.4 to the matrices from Theorem 2.5. □

2.4.5. Perturbations of $\mathcal{L}_m^T \oplus \mathcal{L}_n$

Theorem 2.6. (a) *The set of Kronecker's canonical forms of all pairs in a sufficiently small neighborhood of*

$$\mathcal{L}_m^T \oplus \mathcal{L}_n \tag{2.45}$$

consists of the pairs (2.45) *and*

$$\mathcal{D}_{r_1}(\lambda_1) \oplus \cdots \oplus \mathcal{D}_{r_t}(\lambda_t), \qquad r_1 + \cdots + r_t = m + n - 1, \tag{2.46}$$

with distinct eigenvalues $\lambda_1, \ldots, \lambda_t \in \mathbb{C} \cup \infty$.

(b) *Each pair* (2.46) *with distinct eigenvalues* $\lambda_1, \dots, \lambda_t \in \mathbb{C} \cup \infty$ *is equivalent to a pair of the form*

$$\left(\left[\begin{array}{ccc|ccc} 1 & & & & & \alpha_1 \\ 0 & \ddots & & & & \alpha_2 \\ & \ddots & 1 & & 0 & \vdots \\ & & 0 & & & \alpha_m \\ \hline & & & 1 & 0 & \\ & 0 & & & \ddots & \ddots \\ & & & & 1 & 0 \end{array} \right], \left[\begin{array}{ccc|cccc} 0 & & & \beta_1 & \beta_2 & \dots & \beta_n \\ 1 & \ddots & & & & & \\ & \ddots & 0 & & & 0 & \\ & & 1 & & & & \\ \hline & & & 0 & 1 & & \\ & 0 & & & \ddots & \ddots & \\ & & & & & 0 & 1 \end{array} \right] \right) \tag{2.47}$$

(which is obtained by an arbitrarily small perturbation of (2.45)*), in which*

$$(-\beta_1, \dots, -\beta_n, \alpha_1, \dots, \alpha_m) := \varepsilon(c_0, \dots, c_{r-1}, 1, 0, \dots, 0), \tag{2.48}$$

ε *is an arbitrary nonzero complex number, and* $c_0, \dots, c_{r-1}$ *are defined by*

$$c_0 + c_1 x + \dots + c_{r-1} x^{r-1} + x^r := \prod_{\lambda_i \neq \infty} (x - \lambda_i)^{r_i}.$$

Proof. Let $(C, D) = \mathcal{P}^{\alpha_1 \dots \alpha_m}_{\beta_1 \dots \beta_n}$ denote the pair (2.47). Then

$$\begin{aligned} (D^T, C^T) &= \left(\left[\begin{array}{ccc} c|cR_m & & 0 \\ \hline \beta_1 & & \\ \vdots & 0 & R_n^T \\ \beta_n & & \end{array} \right], \left[\begin{array}{ccc} c|cL_m & & 0 \\ \hline & & \\ & 0 & L_n^T \\ \alpha_1 & \dots & \alpha_m \end{array} \right] \right) \\ &\sim \left(\left[\begin{array}{ccc} & \beta_1 & \\ c|cR_n^T & \vdots & 0 \\ & \beta_n & \\ \hline 0 & & R_m \end{array} \right], \left[\begin{array}{cccc} c|cL_n^T & & 0 & \\ & \alpha_1 & \dots & \alpha_m \\ \hline 0 & & L_m & \end{array} \right] \right) \qquad (2.49) \\ &\sim \left(\left[\begin{array}{ccc} & & \beta_n \\ c|cL_n^T & 0 & \vdots \\ & & \beta_1 \\ \hline 0 & L_m & \end{array} \right], \left[\begin{array}{cccc} & \alpha_m & \dots & \alpha_1 \\ c|cR_n^T & & & \\ & & 0 & \\ \hline 0 & & R_m & \end{array} \right] \right) = \mathcal{P}^{\beta_n \dots \beta_1}_{\alpha_m \dots \alpha_1}; \end{aligned}$$

the third pair is obtained from the second by reversing the order of rows in each horizontal strip and reversing the order of columns in each vertical strip.

By Theorem 1.3, there is a neighborhood of (2.45), in which each pair is equivalent to the pair

$$\mathcal{P}^{\alpha_1\dots\alpha_m}_{\beta_1\dots\beta_n} \tag{2.50}$$

for some $\alpha_1,\dots,\alpha_m,\beta_1,\dots,\beta_n$. The following three cases are possible.

Case 1: $\alpha_m \neq 0$ in (2.50). In this case, $\mathcal{P}^{\alpha_1\dots\alpha_m}_{\beta_1\dots\beta_n} \sim (I_{m+n-1},\Phi)$ with

$$\Phi := \begin{bmatrix} -c_{m+n-2} & \dots & -c_1 & -c_0 \\ 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \end{bmatrix}, \tag{2.51}$$

$$(c_0\dots,c_{m+n-2}) := \alpha_m^{-1}(-\beta_1,\dots,-\beta_n,\alpha_1,\dots,\alpha_{m-1}) \tag{2.52}$$

because

$$\mathcal{P}^{\alpha_1\dots\alpha_m}_{\beta_1\dots\beta_n}(Q_{m-1}\oplus Z_n) = (Q_m\oplus Z_{n-1})(I_{m+n-1},\Phi), \tag{2.53}$$

in which

$$Q_p := \begin{bmatrix} \alpha_m & \alpha_{m-1} & \alpha_{m-2} & \ddots \\ & \alpha_m & \alpha_{m-1} & \ddots \\ & & \alpha_m & \ddots \\ & & & \ddots \end{bmatrix} \ (p\text{-by-}p), \qquad Z_p := \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix} \ (p\text{-by-}p).$$

For example, if $m = n = 4$, then (2.53) takes the form

$$\left(\left[\begin{array}{ccc|cccc} 1&0&0&&&&\alpha_1 \\ 0&1&0&&&&\alpha_2 \\ 0&0&1&&&&\alpha_3 \\ 0&0&0&&&&\alpha_4 \\ \hline &&&1&0&0&0 \\ &&&0&1&0&0 \\ &&&0&0&1&0 \end{array}\right], \left[\begin{array}{ccc|cccc} 0&0&0&\beta_1&\beta_2&\beta_3&\beta_4 \\ 1&0&0&&&& \\ 0&1&0&&&& \\ 0&0&1&&&& \\ \hline &&&0&1&0&0 \\ &&&0&0&1&0 \\ &&&0&0&0&1 \end{array}\right]\right) \left[\begin{array}{ccc|cccc} \alpha_4&\alpha_3&\alpha_2&&&& \\ &\alpha_4&\alpha_3&&&& \\ &&\alpha_4&&&& \\ \hline &&&&&&1 \\ &&&&&1& \\ &&&&1&& \\ &&&1&&& \end{array}\right]$$

$$= \left[\begin{array}{cccc|ccc} \alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 & & & \\ & \alpha_4 & \alpha_3 & \alpha_2 & & & \\ & & \alpha_4 & \alpha_3 & & & \\ & & & \alpha_4 & & & \\ \hline & & & & & & 1 \\ & & & & & 1 & \\ & & & & 1 & & \end{array}\right] \left(I_7, \left[\begin{array}{ccc|cccc} -c_6 & -c_5 & -c_4 & -c_3 & -c_2 & -c_1 & -c_0 \\ 1 & 0 & 0 & & & & \\ 0 & 1 & 0 & & & & \\ 0 & 0 & 1 & & & & \\ \hline & & & 1 & 0 & 0 & 0 \\ & & & 0 & 1 & 0 & 0 \\ & & & 0 & 0 & 1 & 0 \end{array}\right] \right)$$

in which

$$\alpha_4(c_0, c_1, c_2, c_3, c_4, c_5, c_6) = (-\beta_1, -\beta_2, -\beta_3, -\beta_4, \alpha_1, \alpha_2, \alpha_3).$$

The matrix (2.51) has the Jordan canonical form $J_{r_1}(\lambda_1) \oplus \dots \oplus J_{r_t}(\lambda_t)$ with distinct $\lambda_1, \dots, \lambda_t \in \mathbb{C}$; its characteristic polynomial is

$$\begin{aligned} &(x-\lambda_1)^{r_1} \cdots (x-\lambda_t)^{r_t} = c_0 + c_1 x + \dots + c_{m+n-2} x^{m+n-2} + x^{m+n-1} \\ &= \alpha_m^{-1}(-\beta_1 - \beta_2 x - \dots - \beta_n x^{n-1} + \alpha_1 x^n + \alpha_2 x^{n+1} + \dots + \alpha_m x^{m+n-1}). \end{aligned} \tag{2.54}$$

Thus,

$$\mathcal{P}^{\alpha_1 \dots \alpha_m}_{\beta_1 \dots \beta_n} \sim (I, \Phi) \sim \mathcal{D}_{r_1}(\lambda_1) \oplus \dots \oplus \mathcal{D}_{r_t}(\lambda_t) \qquad \text{if } \alpha_m \neq 0, \tag{2.55}$$

which is a pair of the form (2.46). We have proved the statement (a) in Case 1.

By (2.55), each pair (2.46) with distinct *nonzero* eigenvalues $\lambda_1, \dots, \lambda_t \in \mathbb{C} \cup \infty$ is equivalent to $\mathcal{P}^{\alpha_1 \dots \alpha_m}_{\beta_1 \dots \beta_n}$ defined by (2.54). Then (2.52) holds, and so $\mathcal{P}^{\alpha_1 \dots \alpha_m}_{\beta_1 \dots \beta_n}$ is the pair (2.47) defined by (2.48) with $\varepsilon = \alpha_m$. The pair $\mathcal{P}^{\alpha_1 \dots \alpha_m}_{\beta_1 \dots \beta_n}$ is also equivalent to the pair (2.47) defined by (2.48) with an arbitrary nonzero ε since

$$\begin{aligned} &\left(\begin{bmatrix} L_m^T & P \\ 0 & L_n \end{bmatrix}, \begin{bmatrix} R_m^T & Q \\ 0 & R_n \end{bmatrix} \right) \begin{bmatrix} I_{m-1} & 0 \\ 0 & \delta I_n \end{bmatrix} \\ &\qquad = \begin{bmatrix} I_m & 0 \\ 0 & \delta I_{n-1} \end{bmatrix} \left(\begin{bmatrix} L_m^T & \delta P \\ 0 & L_n \end{bmatrix}, \begin{bmatrix} R_m^T & \delta Q \\ 0 & R_n \end{bmatrix} \right) \end{aligned} \tag{2.56}$$

for an arbitrary nonzero δ. This proves the statement (b) if all $\lambda_i \neq \infty$.

Case 2: $\alpha_k \neq 0 = \alpha_{k+1} = \cdots = \alpha_m$ *for some* $k < m$ *in* (2.50). Let us show that

$$\mathcal{P}^{\alpha_1 \dots \alpha_m}_{\beta_1 \dots \beta_n} = \mathcal{P}^{\alpha_1 \dots \alpha_k\, 0 \dots 0}_{\beta_1 \ \dots \ \beta_n} \sim \mathcal{P}^{\alpha_1 \dots \alpha_k}_{\beta_1 \dots \beta_n} \oplus (J_{m-k}(0), I_{m-k}) \qquad \text{if } \alpha_k \neq 0. \tag{2.57}$$

For clarity, we first prove (2.57) in the following special case:

$$\mathcal{P}^{\alpha_1\,\alpha_2\,0\,0}_{\beta_1\beta_2\beta_3\beta_4} \sim \mathcal{P}^{\alpha_1\,\alpha_2}_{\beta_1\beta_2\beta_3\beta_4} \oplus (J_2(0), I_2) \qquad \text{if } \alpha_2 \neq 0. \tag{2.58}$$

The first pair in (2.58) is

$$(C, D) := \left(\left[\begin{array}{ccc|cccc} 1 & 0 & 0 & & & & \alpha_1 \\ 0 & 1 & 0 & & & & \alpha_2 \\ 0 & 0 & 1 & & & & 0 \\ 0 & 0 & 0 & & & & 0 \\ \hline & & & 1 & 0 & 0 & 0 \\ & & & 0 & 1 & 0 & 0 \\ & & & 0 & 0 & 1 & 0 \end{array}\right], \left[\begin{array}{ccc|cccc} 0 & 0 & 0 & \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ 1 & 0 & 0 & & & & \\ 0 & 1 & 0 & & & & \\ 0 & 0 & 1 & & & & \\ \hline & & & 0 & 1 & 0 & 0 \\ & & & 0 & 0 & 1 & 0 \\ & & & 0 & 0 & 0 & 1 \end{array}\right] \right), \quad \alpha_2 \neq 0.$$

It is sufficient to make zero the entry $(2, 2)$ of C; i.e., to prove that

$$(C, D) \sim \left(\left[\begin{array}{ccc|cccc} 1 & 0 & 0 & & & & \alpha_1 \\ 0 & 0 & 0 & & & & \alpha_2 \\ 0 & 0 & 1 & & & & 0 \\ 0 & 0 & 0 & & & & 0 \\ \hline & & & 1 & 0 & 0 & 0 \\ & & & 0 & 1 & 0 & 0 \\ & & & 0 & 0 & 1 & 0 \end{array}\right], \left[\begin{array}{ccc|cccc} 0 & 0 & 0 & \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ 1 & 0 & 0 & & & & \\ 0 & 1 & 0 & & & & \\ 0 & 0 & 1 & & & & \\ \hline & & & 0 & 1 & 0 & 0 \\ & & & 0 & 0 & 1 & 0 \\ & & & 0 & 0 & 0 & 1 \end{array}\right] \right) \tag{2.59}$$

(then the pair $\left(\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]\right)$ in the squares is a direct summand). We make this zero preserving the other entries by the following sequence of elementary transformations with (C, D):

- Substituting column 7 multiplied by α_2^{-1} from column 2, we make zero the entry $(2,2)$ of C:

$$\left(\left[\begin{array}{ccc|cccc} 1 & * & 0 & & & & \alpha_1 \\ 0 & 0 & 0 & & & & \alpha_2 \\ 0 & 0 & 1 & & & & 0 \\ 0 & 0 & 0 & & & & 0 \\ \hline & & & 1 & 0 & 0 & 0 \\ & & & 0 & 1 & 0 & 0 \\ & & & 0 & 0 & 1 & 0 \end{array}\right], \left[\begin{array}{ccc|cccc} 0 & * & 0 & \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ 1 & 0 & 0 & & & & \\ 0 & 1 & 0 & & & & \\ 0 & 0 & 1 & & & & \\ \hline & & & 0 & 1 & 0 & 0 \\ & & & 0 & 0 & 1 & 0 \\ & * & & 0 & 0 & 0 & 1 \end{array}\right]\right).$$

This transformation may spoil the entries denoted by $*$ in columns 2 of C and D; we restore them as follow.

- We restore column 2 of C by adding column 1 (multiplied by a scalar) to column 2. This transformation spoils entry $(2,2)$ of D; we restore it and the entries denoted by stars in column 2 of D by adding row 3 to rows 1, 2, and 7. We obtain

$$\left(\left[\begin{array}{ccc|cccc} 1 & 0 & * & & & & \alpha_1 \\ 0 & 0 & * & & & & \alpha_2 \\ 0 & 0 & 1 & & & & 0 \\ 0 & 0 & 0 & & & & 0 \\ \hline & & & 1 & 0 & 0 & 0 \\ & & & 0 & 1 & 0 & 0 \\ & & * & 0 & 0 & 1 & 0 \end{array}\right], \left[\begin{array}{ccc|cccc} 0 & 0 & 0 & \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ 1 & 0 & 0 & & & & \\ 0 & 1 & 0 & & & & \\ 0 & 0 & 1 & & & & \\ \hline & & & 0 & 1 & 0 & 0 \\ & & & 0 & 0 & 1 & 0 \\ & & & 0 & 0 & 0 & 1 \end{array}\right]\right).$$

- We restore column 3 of C by adding columns 1, 6, and 7, which spoils column 3 of D. We restore it by adding row 4 and obtain (2.59), which proves (2.58).

The equivalence (2.57) for an arbitrary pair $(C,D) = \mathcal{P}^{\alpha_1 \dots \alpha_m}_{\beta_1 \dots \beta_n}$ with $\alpha_k \neq 0 = \alpha_{k+1} = \dots = \alpha_m$ is proved in the same way: we make zero the entry (k,k) of C by adding the last column, which may spoil the entries $(1,k), \dots, (k-$

$1, k)$ of C; they are made zero by adding columns $1, \ldots, k-1$. This spoils column k of D; we restore it by row transformations. This spoils column $k+1$ of C; we restore it by column transformations, and so on, until we obtain the equivalence (2.57).

By (2.55) and (2.57),

$$\mathcal{P}^{\alpha_1 \ldots \alpha_m}_{\beta_1 \ldots \beta_n} = \mathcal{P}^{\alpha_1 \ldots \alpha_k \, 0 \ldots 0}_{\beta_1 \;\ldots\; \beta_n} \sim \mathcal{D}_{r_1}(\lambda_1) \oplus \cdots \oplus \mathcal{D}_{r_{t-1}}(\lambda_{t-1}) \oplus \mathcal{D}_{m-k}(\infty), \tag{2.60}$$

which proves the statement (a) in Case 2. Since

$$\alpha_k^{-1}(-\beta_1 - \cdots - \beta_n x^{n-1} + \alpha_1 x^n + \cdots + \alpha_k x^{n+k-1}) = \prod_{i=1}^{t-1} (x - \lambda_i)^{r_i}, \tag{2.61}$$

the statement (b) holds for $\varepsilon = \alpha_k$. It holds for an arbitrary nonzero ε due to (2.56).

Case 3: $\alpha_1 = \cdots = \alpha_m = 0$ in (2.50); that is, $(C, D) = \mathcal{P}^{0 \;\cdots\; 0}_{\beta_1 \ldots \beta_n}$. If $\beta_1 = \cdots = \beta_n = 0$, then (C, D) is the pair (2.45). Let $\beta_k \neq 0 = \beta_{k-1} = \cdots = \beta_1$ for some $k \geqslant 1$. By (2.49), (2.60), and (2.61), we have

$$(D^T, C^T) \sim \mathcal{P}^{\beta_n \ldots \beta_1}_{0 \;\ldots\; 0} = \mathcal{P}^{\beta_n \ldots \beta_k \, 0 \ldots 0}_{0 \;\ldots\; 0} \sim \mathcal{D}_{r_1}(\mu_1) \oplus \cdots \oplus \mathcal{D}_{r_t}(\mu_t) \oplus \mathcal{D}_{k-1}(\infty),$$

in which $\mu_1, \ldots, \mu_t$ are distinct and

$$(x - \mu_1)^{r_1} \cdots (x - \mu_t)^{r_t} = \beta_k^{-1}(\beta_n x^m + \beta_{n-1} x^{m+1} + \cdots + \beta_k x^{m+n-k}).$$

Let $\beta_n = \beta_{n-1} = \cdots = \beta_{l+1} = 0 \neq \beta_l$ for some $l \geqslant k$. Then

$$(x - \mu_1)^{r_1} \cdots (x - \mu_t)^{r_t} = \beta_k^{-1}(\beta_l x^{m+n-l} + \beta_{l-1} x^{m+n-l+1} + \cdots + \beta_k x^{m+n-k}).$$

Set $\mu_t = 0$ and rewrite this equality as follows:

$$(x - \mu_1)^{r_1} \cdots (x - \mu_{t-1})^{r_{t-1}} x^{m+n-l} = \beta_k^{-1}(\beta_l + \beta_{l-1} x + \cdots + \beta_k x^{l-k}) x^{m+n-l}. \tag{2.62}$$

Therefore,

$$(D^T, C^T) \sim \mathcal{D}_{r_1}(\mu_1) \oplus \cdots \oplus \mathcal{D}_{r_{t-1}}(\mu_{t-1}) \oplus \mathcal{D}_{m+n-l}(0) \oplus \mathcal{D}_{k-1}(\infty),$$

If $k = l$, then $t = 1$ and $(C, D) \sim \mathcal{D}_{m+n-l}(\infty) \oplus \mathcal{D}_{k-1}(0)$.

Let $l > k$, then $t \geqslant 2$. Setting

$$\lambda_1 := \mu_1^{-1},\ \ldots,\ \lambda_{t-1} := \mu_{t-1}^{-1},$$

we find that

$$(C, D) \sim \mathcal{D}_{r_1}(\lambda_1) \oplus \cdots \oplus \mathcal{D}_{r_{t-1}}(\lambda_{t-1}) \oplus \mathcal{D}_{m+n-l}(\infty) \oplus \mathcal{D}_{k-1}(0).$$

This proves the statement (a) in Case 3.

Replacing x by x^{-1} in the polynomials (2.62) and equating the leading coefficients, we obtain

$$(x^{-1} - \lambda_1^{-1})^{r_1} \cdots (x^{-1} - \lambda_{t-1}^{-1})^{r_{t-1}} = \beta_k^{-1}(\beta_l + \beta_{l-1}x^{-1} + \cdots + \beta_k x^{k-l}),$$

and so

$$(x - \lambda_1)^{r_1} \cdots (x - \lambda_{t-1})^{r_{t-1}} = \beta_l^{-1}(\beta_k + \beta_{k+1}x + \cdots + \beta_l x^{k-l}).$$

This proves the statement (b) for $\varepsilon = -\beta_l$. It holds for an arbitrary nonzero ε due to (2.56). $\square$

2.4.6. Perturbations of $\mathcal{D}_m(\lambda) \oplus \mathcal{D}_n(\lambda)$

Theorem 2.7. (a) *If a Kronecker pair $\mathcal{K}$ is equivalent to a pair in an arbitrarily small neighborhood of*

$$\mathcal{D}_m(\lambda) \oplus \mathcal{D}_n(\lambda), \qquad m \leqslant n, \tag{2.63}$$

then $\mathcal{K}$ has the form

$$\mathcal{D}_{m-r}(\lambda) \oplus \mathcal{D}_{n+r}(\lambda), \qquad 0 \leqslant r \leqslant m. \tag{2.64}$$

(b) *Each pair (2.64) with $r > 0$ is equivalent to a pair of the form*

$$\begin{array}{ll} \left(I_{m+n}, \begin{bmatrix} J_m(\lambda) & \Delta_r(\varepsilon)^T \\ 0 & J_n(\lambda) \end{bmatrix}\right) & \textit{if } \lambda \in \mathbb{C} \\ \left(\begin{bmatrix} J_m(0) & \Delta_r(\varepsilon)^T \\ 0 & J_n(0) \end{bmatrix}, I_{m+n}\right) & \textit{if } \lambda = \infty \end{array}$$

(which is obtained by an arbitrarily small perturbation of (2.63)*), in which ε is an arbitrary nonzero complex number.*

Proof. This theorem follows from Theorem 2.9 by the reasons that are given at the beginning of the next section. □

2.5. Perturbations of Jordan matrices

Each matrix that is obtained by an arbitrarily small perturbation of I_n is reduced to I_n by equivalence transformations that are close to the identity transformation. Hence, each pair that is obtained by an arbitrarily small perturbation of (I_n, B) is reduced to a pair of the form (I_n, C) by equivalence transformations that are close to the identity transformation.

Hence, the theory of perturbations of matrix pairs (A, B) with a nonsingular A under equivalence is reduced to the theory of perturbations of square matrices under similarity. By Theorem 2.1, it reduces to the theory of perturbations of Jordan matrices with a single eigenvalue.

Theorem 2.8. *Let J be a Jordan matrix with a single eigenvalue λ.*

(a) *If J is a Jordan block, then $\langle J\rangle$ has no successors.*

(b) *Let J have at least 2 Jordan blocks. Write it as follows:*

$$J = P \oplus J_p(\lambda) \oplus J_q(\lambda) \oplus Q, \qquad p \leqslant q, \tag{2.65}$$

in which P is a direct sum of Jordan blocks of sizes $\leqslant p$ and Q is a direct sum of Jordan blocks of sizes $\geqslant q$ (P and/or Q can be zero). Define the Jordan matrix

$$J_{p,q} := P \oplus J_{p-1}(\lambda) \oplus J_{q+1}(\lambda) \oplus Q, \tag{2.66}$$

in which $J_{p-1}(\lambda)$ is absent if $p = 1$. Then $\langle J_{p,q}\rangle$ immediately succeeds $\langle J\rangle$, and each immediate successor of $\langle J\rangle$ is $\langle J_{p,q}\rangle$ for some p and q.

The *Weyr characteristic* of a square matrix A for an eigenvalue λ is the non-increasing sequence $(m_1, m_2, \dots)$ in which m_i is the number of Jordan blocks $J_l(\lambda)$ of size $l \geqslant i$ in the Jordan form of A.

Each nilpotent matrix A is similar to a matrix of the form

$$W = \begin{bmatrix} 0_{m_1} & F_1 & & 0 \\ & 0_{m_2} & \ddots & \\ & & \ddots & F_{k-1} \\ 0 & & & 0_{m_k} \end{bmatrix}, \qquad F_i := \begin{bmatrix} I_{m_{i+1}} \\ 0 \end{bmatrix}, \tag{2.67}$$

which was called in [65] the *Weyr canonical form* of A (see historical remarks in [**?**, pp. 80–82]). The Weyr characteristic of A for its single eigenvalue 0 is $(m_1, m_2, \dots)$ since

$$W^2 = \begin{bmatrix} 0_{m_1} & 0 & F_1F_2 & & 0 \\ & 0_{m_2} & 0 & \ddots & \\ & & 0_{m_3} & \ddots & F_{k-2}F_{k-1} \\ & & & \ddots & 0 \\ 0 & & & & 0_{m'_k} \end{bmatrix}, \quad \dots$$

implies

$$\begin{aligned} m_1 &= \operatorname{nullity} W = \operatorname{nullity} A, \\ m_1 + m_2 &= \operatorname{nullity} W^2 = \operatorname{nullity} A^2, \\ m_1 + m_2 + m_3 &= \operatorname{nullity} W^3 = \operatorname{nullity} A^3, \\ &\dots\dots\dots \end{aligned}$$

Hence W is uniquely determined by A. The Weir canonical form (2.67) of A is permutation similar to the Jordan canonical form of A.

Lemma 2.2. *Let J and J' be Jordan matrices with a single eigenvalue λ. Let $(m_1, m_2, \dots)$ and $(m'_1, m'_2, \dots)$ be their Weyr characteristics. Write*

$$s_i := m_1 + \dots + m_i, \qquad s'_i := m'_1 + \dots + m'_i \tag{2.68}$$

for $i = 1, 2, \dots$. Then

$$\langle J \rangle \leqslant \langle J' \rangle \iff s_i \geqslant s'_i \text{ for all } i. \tag{2.69}$$

Example 2.1. Let

$$J = J_3(\lambda) \oplus J_4(\lambda) \oplus J_4(\lambda), \qquad J' = J_3(\lambda) \oplus J_3(\lambda) \oplus J_5(\lambda).$$

Then

$$\begin{array}{llll} m_1 = m_2 = m_3 = 3, & m_4 = 2, & m_5 = 0, & m_6 = m_7 = \cdots = 0, \\ m'_1 = m'_2 = m'_3 = 3, & m'_4 = 1, & m'_5 = 1, & m'_6 = m'_7 = \cdots = 0, \end{array}$$

and so

$$\begin{array}{lllll} s_1 = 3, & s_2 = 6, & s_3 = 9, & s_4 = 11, & s_5 = s_6 = \cdots = 11, \\ s'_1 = 3, & s'_2 = 6, & s'_3 = 9, & s'_4 = 10, & s'_5 = s'_6 = \cdots = 11. \end{array}$$

Hence, $\langle J \rangle < \langle J' \rangle$.

Proof of Lemma 2.2. Let J be a Jordan matrix with a single eigenvalue λ. Then $\langle J - \lambda I \rangle = \langle J \rangle - \lambda I$ and $\overline{\langle J - \lambda I \rangle} = \overline{\langle J \rangle} - \lambda I$ for their closures. Hence, we must prove (2.69) only for J and J' with the single eigenvalue $\lambda = 0$.

$\Longrightarrow$. Let J' be a Jordan matrix such that each neighborhood of J contains a matrix whose Jordan canonical form is J'. This means that there is a convergent sequence

$$A_1, A_2, \ldots \to J \tag{2.70}$$

in which all A_i are similar to J'. All A_i have the same characteristic polynomial $f(x)$. Since the coefficients of the characteristic polynomial continuously depend on the matrix entries, $f(x)$ is also the characteristic polynomial of J. Hence, $f(x) = x^n$, and so J' is nilpotent.

Since all A_i are similar to J', they have the same Weyr canonical form

$$S_i^{-1} A_i S_i = \begin{bmatrix} 0_{m'_1} & F'_1 & & 0 \\ & 0_{m'_2} & \ddots & \\ & & \ddots & F'_{k-1} \\ 0 & & & 0_{m'_k} \end{bmatrix}, \qquad F'_i := \begin{bmatrix} I_{m'_{i+1}} \\ 0 \end{bmatrix},$$

in which $(m_1', m_2', \dots)$ is the Weyr characteristic of J'. Applying the Gram–Schmidt orthogonalisation process to the columns of S_i, we get a unitary matrix $U_i = S_i R_i$, where R_i is a nonsingular upper-triangular matrix. Then

$$U_i^{-1} A_i U_i = R_i^{-1} \cdot S_i^{-1} A_i S_i \cdot R_i = \begin{bmatrix} 0_{m_1'} & V_1^{(i)} & * & \dots & * \\ & 0_{m_2'} & V_2^{(i)} & \ddots & \vdots \\ & & 0_{m_3'} & \ddots & * \\ & & & \ddots & V_{k-1}^{(i)} \\ 0 & & & & 0_{m_k'} \end{bmatrix},$$

in which every $V_j^{(i)}$ is an $m_i' \times m_{i+1}'$ matrix with linearly independent columns.

The set of matrices $U_1, U_2, \dots$ is bounded since each entry of a unitary matrix has modulus $\leqslant 1$. Hence this set has a limit point, which we denote by U. Deleting some A_i in (2.70) if necessarily, we make $U_i \to U$. Since each U_i is unitary, we have $U_i U_i^* = I$, and so $UU^* = I$. Hence U is unitary and

$$U_i^{-1} A_i U_i \to U^{-1} J U = \begin{bmatrix} 0_{m_1'} & V_1 & * & \dots & * \\ & 0_{m_2'} & V_2 & \ddots & \vdots \\ & & 0_{m_3'} & \ddots & * \\ & & & \ddots & V_{k-1} \\ 0 & & & & 0_{m_k'} \end{bmatrix},$$

in which $V_1^{(i)} \to V_1, \dots, V_{k-1}^{(i)} \to V_{k-1}$. Note that the columns of some V_i can be linearly dependent.

Therefore,

$$m_1 = \text{nullity}\, J = \text{nullity}\, U^{-1} J U \geqslant m_1'.$$

Since

$$U^{-1}J^2U = \begin{bmatrix} 0_{m_1'} & 0 & V_1V_2 & & 0 \\ & 0_{m_2'} & 0 & \ddots & \\ & & 0_{m_3'} & \ddots & V_{k-2}V_{k-1} \\ & & & \ddots & 0 \\ 0 & & & & 0_{m_k'} \end{bmatrix},$$

we have

$$m_1 + m_2 = \text{nullity}\, J^2 = \text{nullity}\, U^{-1}J^2U \geqslant m_1' + m_2',$$

and so on, which proves "$\Longrightarrow$" in (2.69).

$\Longleftarrow$. Let W and W' be Weyr canonical matrices of the same size with Weyr characteristics $(m_1, m_2, \dots, m_k)$ and $(m_1', m_2', \dots, m_k')$ satisfying $s_1 \geqslant s_1'$, $s_2 \geqslant s_2'$, ... These inequalities guarantee that for each sufficiently small ε the Weyr canonical form of $\varepsilon W' + W$ is W'. If $\varepsilon_i \to 0$, then $\varepsilon_i W' + W \to W$. Hence $\langle W \rangle \leqslant \langle W' \rangle$. □

Proof of Theorem 2.8. (a) Let $J = J_p(\lambda)$, and let $\langle J \rangle \leqslant \langle J' \rangle$. By (2.69), $m_1' \leqslant m_1 = 1$. However, m_1' is the number of Jordan blocks in J'. Hence, J' is a Jordan block. Since J and J' have the same size, $J' = J_p(\lambda) = J$.

(b) Denote by $(m_1(X), m_2(X), \dots)$ the Weyr characteristic of a matrix X and write $s_i(X) := m_1(X) + \cdots + m_i(X)$. Let A, B, and C be square matrices with a single eigenvalue. Since $m_i(A \oplus B) = m_i(A) + m_i(B)$, we have $s_i(A \oplus B) = s_i(A) + s_i(B)$. Thus, $s_i(A \oplus B) \leqslant s_i(A \oplus C)$ if and only if $s_i(B) \leqslant s_i(C)$. By (2.69),

$$\langle A \oplus B \rangle \leqslant \langle A \oplus C \rangle \quad \Longleftrightarrow \quad \langle B \rangle \leqslant \langle C \rangle. \tag{2.71}$$

Let $(m_1, m_2, \dots)$ and $(\widetilde{m}_1, \widetilde{m}_2, \dots)$ be the Weyr characteristics of the matrices (2.65) and (2.66). Then $\widetilde{m}_p = m_p - 1$, $\widetilde{m}_{q+1} = m_{q+1} + 1$, the other $\widetilde{m}_i = m_i$, and so

$$\widetilde{s}_p = s_p - 1,\ \widetilde{s}_{p+1} = s_{p+1} - 1,\ \dots,\ \widetilde{s}_q = s_q - 1, \quad \text{the other } \widetilde{s}_i = s_i \tag{2.72}$$

in the notation (2.68). Let us prove the following three facts.

Fact 1: $\langle J\rangle < \langle J_{p,q}\rangle$. Due to (2.71), this inequality follows from $\langle J_p(\lambda)\oplus J_q(\lambda)\rangle < \langle J_{p-1}(\lambda)\oplus J_{q+1}(\lambda)\rangle$, which holds by (2.69) and (2.72).

Fact 2: if J' is a Jordan matrix with the single eigenvalue λ, then

$$\langle J\rangle < \langle J'\rangle \quad\Longrightarrow\quad \langle J\rangle < \langle J_{p,q}\rangle \leqslant \langle J'\rangle \text{ for some } p,q. \tag{2.73}$$

Due to (2.71), it is sufficient to prove (2.73) for J and J' that have no common Jordan blocks. By the assumptions of Theorem 2.8(b), J has at least two Jordan blocks. Let p and q be such that

$$J = J_p(\lambda)\oplus J_q(\lambda)\oplus Q, \qquad p\leqslant q,$$

in which all Jordan blocks of Q are of size $\geqslant q$. Let us prove that

$$J_{p,q} = J_{p-1}(\lambda)\oplus J_{q+1}(\lambda)\oplus Q$$

satisfies (2.73).

By Lemma 2.2 and $\langle J'\rangle \geqslant \langle J\rangle$, $s'_i \leqslant s_i$ for all i. By Step 1, $\langle J_{p,q}\rangle > \langle J\rangle$. We must prove that $\langle J'\rangle \geqslant \langle J_{p,q}\rangle$; that is, $s'_i \leqslant \widetilde{s}_i$ for all i. Due to (2.72), it suffices to prove that

$$s'_p < s_p, \quad s'_{p+1} < s_{p+1}, \quad \dots, \quad s'_q < s_q. \tag{2.74}$$

Since J and J' do not have common Jordan blocks, J' does not have $J_p(\lambda)$, and so

$$\begin{array}{l} s_1 = m_1 = \cdots = m_p > m_{p+1} \\ \rotatebox{90}{$\leqslant$} \\ s'_1 = m'_1 \geqslant \cdots \geqslant m'_p = m'_{p+1} \end{array}$$

Thus, $m_p \geqslant m'_p$.

If $m_p = m'_p$, then

$$\begin{array}{l} s_1 = m_1 = \cdots = m_p > m_{p+1} \\ \rotatebox{90}{$=$} \\ s'_1 = m'_1 = \cdots = m'_p = m'_{p+1} \end{array}$$

Hence, $s_1 = s'_1$, $s_2 = s'_2$, $\dots$, $s_p = s'_p$, $s_{p+1} = s_p + m_{p+1} < s'_p + m'_{p+1} = s'_{p+1}$, which contradicts $s_{p+1} \geqslant s'_{p+1}$.

Therefore, $m_p > m'_p$, $s_p = s_{p-1} + m_p > s'_{p-1} + m'_p = s'_p$, and so $s_p > s'_p$, which proves (2.74) if $p = q$.

Let $p < q$. Then J has only one $J_p(\lambda)$, which means that $m_p = m_{p+1} + 1$. Since $m_p > m'_p$, we have $m_p - 1 \geqslant m'_p$, and so

$$\begin{array}{ccccc} m_p - 1 = m_{p+1} = m_{p+2} = \dots = m_q \\ \rotatebox{90}{$\leqslant$} \\ m'_p = m'_{p+1} \geqslant m'_{p+2} \geqslant \dots \geqslant m'_q \end{array}$$

We obtain consistently $s_p > s'_p$, $s_{p+1} = s_p + m_{p+1} > s'_p + m'_{p+1} = s'_{p+1}$, $\dots$, $s_q = s_{q-1} + m_q > s'_{q-1} + m'_q = s'_q$, which proves (2.74) if $p < q$.

Fact 3: if J' is a Jordan matrix with the single eigenvalue λ, then

$$\langle J \rangle < \langle J' \rangle \leqslant \langle J_{p,q} \rangle \quad \Longrightarrow \quad J' = J_{p,q}$$

up to permutations of Jordan blocks in J'.

On the contrary, let $\langle J \rangle < \langle J' \rangle < \langle J_{p,q} \rangle$ for some J'. By Fact 2, we can take $J' = J_{p',q'}$ for some $p' \leqslant q'$.

Write $t(J) := (t_1, t_2, \dots)$, in which t_i is the number of $i \times i$ Jordan blocks in J. Then $s(J) := t_1 + t_2 + \dots$ is the number of Jordan blocks in J.

Let $u = (u_1, \dots, u_s)$ and $u = (v_1, \dots, v_s)$ be two sequences of nonnegative integers. Define two lexicographical orders:

$$\begin{aligned} &u \overset{l}{\leq} v \quad \text{if } u = v \text{ or } u_1 = v_1, \dots, u_{k-1} = v_{k-1}, u_k < v_k \text{ for some } k \geqslant 1; \\ &u \overset{r}{\leq} v \quad \text{if } u = v \text{ or } u_k < v_k, u_{k+1} = v_{k+1}, u_{k+2} = v_{k+2}, \dots \text{ for some } k \geqslant 1. \end{aligned}$$

By Fact 2, the inequality $\langle J_{p',q'} \rangle < \langle J_{p,q} \rangle$ implies that $J_{p,q}$ is obtained from $J_{p',q'}$ by a sequence of replacements of type $J \downarrow J_{s,r}$:

$$J_{p',q'} \downarrow (J_{p',q'})_{r_1,s_1} \downarrow ((J_{p',q'})_{r_1,s_1})_{r_2,s_2} \downarrow \cdots \downarrow J_{p,q}. \tag{2.75}$$

Therefore,

(i) $s(J_{p',q'}) \geqslant s(J_{p,q})$,

(ii) if $s(J_{p',q'}) = s(J_{p,q})$, then $t(J_{p',q'}) \overset{l}{\preceq} t(J_{p,q})$, and

(iii) $t(J_{p',q'}) \overset{r}{\succeq} t(J_{p,q})$

since the analogous statements hold for each of the replacements (2.75).

Let $s(J_{p',q'}) > s(J_{p,q})$. Then $J = J_1(\lambda) \oplus \cdots$ and $p = 1$. Hence $q \leqslant p'$, and so $t(J_{p',q'}) \overset{r}{\prec} t(J_{p,q})$, which contradicts (iii).

Thus, $s(J_{p',q'}) = s(J_{p,q})$. If $p' < p$, then (ii) does not hold. If $q' > q$, then (iii) does not hold. Hence, $p \leqslant p' \leqslant q' \leqslant q$, which contradicts with $(p', q') \neq (p, q)$. $\square$

Theorem 2.9. (a) *All matrices in a sufficiently small neighborhood of*

$$J_m(\lambda) \oplus J_n(\lambda), \qquad m \leqslant n$$

are similar to matrices of the form

$$J_{m-r}(\lambda) \oplus J_{n+r}(\lambda), \qquad 0 \leqslant r \leqslant m. \tag{2.76}$$

(b) *Each matrix* (2.76) *with* $r > 0$ *is similar to*

$$\begin{bmatrix} J_m(\lambda) & \Delta_r(\varepsilon)^T \\ 0 & J_n(\lambda) \end{bmatrix}, \tag{2.77}$$

in which $\Delta_r(\varepsilon)$ *with an arbitrary nonzero* $\varepsilon \in \mathbb{C}$ *is defined in* (2.5).

Proof. (a) This statement follows from Theorem 2.8(b).

(b) We make $\varepsilon = 1$ in (2.77) preserving the other entries by the following similarity transformation: we divide by ε the m rows of the first horizontal strip, then multiply by ε the m columns of the first vertical strip. In the obtained matrix

$$
\begin{array}{c|c|c|c|l}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{\scriptstyle r+1 \;\;\; m} & \multicolumn{1}{c}{\scriptstyle 1 \;\;\; m-r} & \multicolumn{1}{c}{} & \\
\begin{matrix}\lambda & 1 & \\ & \ddots & \ddots \\ & & \lambda\end{matrix} & \begin{matrix} \\ \\ 1 \end{matrix} & \begin{matrix}\vdots \\ 0 \\ 1\end{matrix} & & \begin{matrix}\scriptstyle 1 \\ \\ \scriptstyle r\end{matrix} \\
\hline
 & \boxed{\begin{matrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{matrix}} & \begin{matrix}0\\0\\ \vdots \\ 0\end{matrix} & & \begin{matrix}\scriptstyle r+1 \\ \\ \\ \scriptstyle m\end{matrix}\\
\hline
 & \boxed{\begin{matrix}0 & \cancel{0} & & \\ & 0 & \ddots & \\ & & \ddots & \cancel{0} \\ & & & 0\end{matrix}} & \boxed{\begin{matrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{matrix}} & \begin{matrix}\\ \\ \\ 1\end{matrix} & \begin{matrix}\scriptstyle 1 \\ \\ \\ \scriptstyle m-r\end{matrix} \\
\hline
 & & & \begin{matrix}\lambda & \ddots & \\ & \ddots & 1 \\ & & \lambda\end{matrix} & \begin{matrix} \\ \\ \scriptstyle n\end{matrix}
\end{array}
\qquad (2.78)
$$

we make zero the entry "1" to the left of $\varepsilon = 1$ by the following similarity transformations (every $\cancel{0}$ denotes the zero entry that is transformed to -1 and then is restored to 0; compare with (2.31)):

- Make zero the entry "1" to the left of $\varepsilon = 1$ by subtracting the columns $1, 2, \dots, m-r$ of the second vertical strip from the columns $r+1, r+2, \dots, m$ of the first vertical strip, respectively. Thus, the marked $(m-r) \times (m-r)$ subblock in the $(2,2)$th block of the matrix (2.78) is subtracted from the marked $(m-r) \times (m-r)$ subblock in the $(2,1)$th block.

- Make the inverse transformations of rows, adding the rows $r+1, \dots, m$ of the first horizontal strip to the rows $1, \dots, m-r$ of the second horizontal strip. Thus, the $(m-r) \times (m-r)$ subblock in the $(1,1)$th block is added to the $(m-r) \times (m-r)$ subblock in the $(2,1)$th block, restoring it.

The $(m-r) \times (m-r)$ marked subblock in the $(1,1)$th block of the obtained matrix is a direct summand, and so the obtained matrix is permutation similar to (2.76). $\square$

CHAPTER 3

APPLICATIONS OF PERTURBATION THEORY

3.1. Regularizing algorithm for mixed matrix pencils

The text of this section coincides with the text of my paper [51] (up to the numeration of statements, formulas and references).

Van Dooren [70] gave an algorithm that for each pair (A, B) of complex matrices of the same size constructs its *regularizing decomposition*; that is, it constructs a matrix pair that is simultaneously equivalent to (A, B) and has the form

$$(A_1, B_1) \oplus \cdots \oplus (A_t, B_t) \oplus (\underline{A}, \underline{B})$$

in which $(\underline{A}, \underline{B})$ is a pair of nonsingular matrices and each other summand has one of the forms:

$$(F_n, G_n), \quad (F_n^T, G_n^T), \quad (I_n, J_n(0)), \quad (J_n(0), I_n),$$

where $J_n(0)$ is the singular Jordan block and

$$F_n := \begin{bmatrix} 0 & & 0 \\ 1 & \ddots & \\ & \ddots & 0 \\ 0 & & 1 \end{bmatrix}, \quad G_n := \begin{bmatrix} 1 & & 0 \\ 0 & \ddots & \\ & \ddots & 1 \\ 0 & & 0 \end{bmatrix}$$

are $n \times (n-1)$ matrices; $n \geqslant 1$. Note that $(F_1, G_1) = (0_{10}, 0_{10})$; we denote by 0_{mn} the zero matrix of size $m \times n$, where $m, n \in \{0, 1, 2, \dots\}$. The algorithm uses only unitary transformations, which improves its computational stability.

We extend Van Dooren's algorithm to square complex matrices up to consimilarity transformations $A \mapsto SA\bar{S}^{-1}$ and to pairs of $m \times n$ matrices up to transformations $(A, B) \mapsto (SAR, SB\bar{R})$, in which S and R are nonsingular matrices.

A regularizing algorithm for matrices of undirected cycles of linear mappings was constructed by Sergeichuk [65] and, independently, by Varga [71]. A regularizing algorithm for matrices under congruence was constructed by Horn and Sergeichuk [46].

All matrices that we consider are complex matrices.

3.1.1. Regularizing unitary algorithm for matrices under consimilarity

Two matrices A and B are *consimilar* if there exists a nonsingular matrix S such that $SA\bar{S}^{-1} = B$. Two matrices are consimilar if and only if they give the same semilinear operator, but in different bases. Recall that a mapping $\mathcal{A} : U \to V$ between complex vector spaces is *semilinear* if

$$\mathcal{A}(au_1 + bu_2) = \bar{a}\mathcal{A}u_1 + \bar{b}\mathcal{A}u_2$$

for all $a, b \in \mathbb{C}$ and $u_1, u_2 \in U$.

The canonical form of a matrix under consimilarity is the following (see [43] or [44]):

Each square complex matrix is consimilar to a direct sum, uniquely determined up to permutation of direct summands, of matrices of the following types:

- *a Jordan block* $J_k(\lambda)$ *with* $\lambda \geqslant 0$*, and*

- $\begin{bmatrix} 0 & 1 \\ \mu & 0 \end{bmatrix}$ *with* $\mu \notin \mathbb{R}$ *or* $\mu < 0$.

Thus, each square matrix A is consimilar to a direct sum

$$J_{n_1}(0) \oplus \cdots \oplus J_{n_k}(0) \oplus \underline{A},$$

in which $\underline{A}$ is nonsingular and is determined up to consimilarity; the other summands are uniquely determined up to permutation. This sum is called a *regularizing decomposition* of A. The following algorithm admits to construct a regularizing decomposition using only unitary transformations.

Algorithm 3.1. *Let A be a singular $n \times n$ matrix. By unitary transformations of rows, we reduce it to the form*

$$S_1 A = \begin{bmatrix} 0_{r_1 n} \\ A' \end{bmatrix}, \quad S_1 \text{ is unitary,}$$

in which the rows of A' are linearly independent. Then we make the coninverse transformations of columns and obtain

$$S_1 A \bar{S}_1^{-1} = \begin{bmatrix} 0_{r_1} & 0 \\ \star & A_1 \end{bmatrix}$$

We apply the same procedure to A_1 and obtain

$$S_2 A_1 \bar{S}_2^{-1} = \begin{bmatrix} 0_{r_2} & 0 \\ \star & A_2 \end{bmatrix}, \quad S_2 \text{ is unitary,}$$

in which the rows of $[\star \ A_2]$ are linearly independent.

We repeat this procedure until we obtain

$$S_t A_{t-1} \bar{S}_t^{-1} = \begin{bmatrix} 0_{r_t} & 0 \\ \star & A_t \end{bmatrix}, \quad S_t \text{ is unitary,}$$

in which A_t is nonsingular. The result of the algorithm is the sequence $r_1, r_2, \ldots, r_t, A_t$.

For a matrix A and a nonnegative integer n, we write

$$A^{(n)} := \begin{cases} 0_{00}, & \text{if } n = 0; \\ A \oplus \dots \oplus A \ (n \text{ summands}), & \text{if } n \geqslant 1. \end{cases}$$

Theorem 3.2. *Let $r_1, r_2, \dots, r_t, A_t$ be obtained by applying Algorithm 1 to a square complex matrix A. Then*

$$r_1 \geqslant r_2 \geqslant \ldots \geqslant r_t \tag{3.1}$$

and A is consimilar to

$$J_1{}^{(r_1-r_2)} \oplus J_2{}^{(r_2-r_3)} \oplus \dots \oplus J_{t-1}^{(r_{t-1}-r_t)} \oplus J_t^{(r_t)} \oplus A_t \tag{3.2}$$

in which $J_k := J_k(0)$ and A_t is determined by A up to consimilarity and the other summands are uniquely determined.

Proof. Let $\mathcal{A} : V \to V$ be a semilinear operator whose matrix in some basis is A. Let $W := \mathcal{A}V$ be the image of $\mathcal{A}$. Then the matrix of the restriction $\mathcal{A}_1 : W \to W$ of $\mathcal{A}$ on W is A_1. Applying Algorithm 1 to A_1, we get the sequence $r_2, \dots, r_t, A_t$. Reasoning by induction on the length t of the algorithm, we suppose that $r_2 \geqslant r_3 \geqslant \ldots \geqslant r_t$ and that A_1 is consimilar to

$$J_1{}^{(r_2-r_3)} \oplus \dots \oplus J_{t-2}^{(r_{t-1}-r_t)} \oplus J_{t-1}^{(r_t)} \oplus A_t. \tag{3.3}$$

Thus, $\mathcal{A}_1 : W \to W$ is given by the matrix (3.3) in some basis of W.

The direct sum (3.3) defines the decomposition of W into the direct sum of invariant subspaces

$$W = (W_{21} \oplus \dots \oplus W_{2,r_2-r_3}) \oplus \dots \oplus (W_{t1} \oplus \dots \oplus W_{tr_t}) \oplus W'.$$

Each W_{pq} is generated by some basis vectors $e_{pq2}, e_{pq3}, \dots, e_{pqp}$ such that

$$\mathcal{A} : \ e_{pq2} \mapsto e_{pq3} \mapsto \dots \mapsto e_{pqp} \mapsto 0.$$

For each W_{pq}, we choose $e_{pq1} \in V$ such that $\mathcal{A}e_{pq1} = e_{pq2}$. The set

$$\{e_{pqp} \,|\, 2 \leqslant p \leqslant t, \ 1 \leqslant q \leqslant r_p - r_{p+1}\} \quad (r_{t+1} := 0)$$

consists of r_2 basis vectors belonging to the kernel of $\mathcal{A}$; we supplement this set to a basis of the kernel of $\mathcal{A}$ by some vectors $e_{111}, \dots, e_{1,r_1-r_2,1}$.

The set of vectors e_{pqs} supplemented by the vectors of some basis of W' is a basis of V. The matrix of $\mathcal{A}$ in this basis has the form (3.2) because

$$\mathcal{A}\colon\ e_{pq1} \mapsto e_{pq2} \mapsto e_{pq3} \mapsto \cdots \mapsto e_{pqp} \mapsto 0$$

for all $p = 1, \dots, t$ and $q = 1, \dots, r_p - r_{p+1}$. This completes the proof of Theorem 3.2.

Example 3.1. Let a square matrix A define a semilinear operator $\mathcal{A}: V \to V$ and let the singular part of its regularizing decomposition be $J_2 \oplus J_3 \oplus J_4$. This means that V possesses a set of linear independent vectors forming the Jordan chains

$$\begin{aligned} \mathcal{A}\colon\quad & e_1 \mapsto e_2 \mapsto e_3 \mapsto e_4 \mapsto 0 \\ & f_1 \mapsto f_2 \mapsto f_3 \mapsto 0 \\ & g_1 \mapsto g_2 \mapsto 0 \end{aligned} \tag{3.4}$$

Applying the first step of Algorithm 1, we get A_1 whose singular part corresponds to the chains

$$\begin{aligned} \mathcal{A}\colon\quad & e_2 \mapsto e_3 \mapsto e_4 \mapsto 0 \\ & f_2 \mapsto f_3 \mapsto 0 \\ & g_2 \mapsto 0 \end{aligned}$$

On the second step, we delete e_2, f_2, g_2 and so on. Thus, r_i is the number of vectors in the ith column of (3.32): $r_1 = 3,\ r_2 = 3,\ r_3 = 2, r_4 = 1$. We get the singular part of regularizing decomposition of A:

$$J_1{}^{(r_1-r_2)} \oplus \cdots \oplus J_{t-1}^{(r_{t-1}-r_t)} \oplus J_t^{(r_t)} = J_1{}^{(3-3)} \oplus J_2{}^{(3-2)} \oplus J_3^{(2-1)} \oplus J_4^{(1)} = J_2 \oplus J_3 \oplus J_4.$$

In particular, if

$$A = \begin{array}{cc} \left[\begin{array}{cccc|ccc|cc} 0 & 0 & 0 & 0 & & & & & \\ 1 & 0 & 0 & 0 & & & & & \\ 0 & 1 & 0 & 0 & & & & & \\ 0 & 0 & 1 & 0 & & & & & \\ \hline & & & & 0 & 0 & 0 & & \\ & & & & 1 & 0 & 0 & & \\ & & & & 0 & 1 & 0 & & \\ \hline & & & & & & & 0 & 0 \\ & & & & & & & 1 & 0 \end{array}\right] & \begin{array}{c} e_1 \\ e_2 \\ e_3 \\ e_4 \\ f_1 \\ f_2 \\ f_3 \\ g_1 \\ g_2 \end{array} \\ \begin{array}{ccccccccc} e_1 & e_2 & e_3 & e_4 & f_1 & f_2 & f_3 & g_1 & g_2 \end{array} & \end{array}, \tag{3.5}$$

then we can apply Algorithm 1 using only transformations of permutational similarity and obtain

$$\begin{array}{cc} \left[\begin{array}{ccc|ccc|cc|c} 0 & 0 & 0 & & & & & & \\ 0 & 0 & 0 & & & & & & \\ 0 & 0 & 0 & & & & & & \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & & & \\ 0 & 1 & 0 & 0 & 0 & 0 & & & \\ 0 & 0 & 1 & 0 & 0 & 0 & & & \\ \hline & & & 1 & 0 & 0 & 0 & 0 & \\ & & & 0 & 1 & 0 & 0 & 0 & \\ \hline & & & & & & 1 & 0 & 0 \end{array}\right] & \begin{array}{c} e_1 \\ f_1 \\ g_1 \\ e_2 \\ f_2 \\ g_2 \\ e_3 \\ f_3 \\ e_4 \end{array} \\ \begin{array}{ccccccccc} e_1 & f_1 & g_1 & e_2 & f_2 & g_2 & e_3 & f_3 & e_4 \end{array} & \end{array}$$

(all unspecified blocks are zero), which is the Weyr canonical form of (3.5), see [44].

3.1.2. Regularizing unitary algorithm for matrix pairs under mixed equivalence

We say that pairs of $m \times n$ matrices (A, B) and (A', B') are *mixed equivalent* if there exist nonsingular S and R such that

$$(SAR, SB\bar{R}) = (A', B').$$

The *direct sum* of matrix pairs (A, B) and (C, D) is defined as follows:

$$(A, B) \oplus (C, D) = \left(\begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix}, \begin{bmatrix} B & 0 \\ 0 & D \end{bmatrix} \right).$$

The canonical form of a matrix pair under mixed equivalence was obtained by Djoković [17] (his result was extended to undirected cycles of linear and semilinear mappings in [15]):

Each pair (A, B) of matrices of the same size is mixed equivalent to a direct sum, determined uniquely up to permutation of summands, of pairs of the following types:

$$(I_n, J_n(\lambda)),\ (I_{2k}, H_{2k}(\mu)),\ (J_n(0), I_n),\ (F_n, G_n),\ (F_n^T, G_n^T),$$

in which λ is real nonnegative number,

$$H_{2k}(\mu) := \begin{bmatrix} 0 & I_k \\ J_k(\mu) & 0 \end{bmatrix},$$

and $\mu \notin \mathbb{R}$ or $\mu < 0$.

Thus, (A, B) is mixed equivalent to a direct sum of a pair $(\underline{A}, \underline{B})$ of nonsingular matrices and summands of the types:

$$(I_n, J_n(0)),\ (J_n(0), I_n),\ (F_n, G_n),\ (F_n^T, G_n^T),$$

in which $(\underline{A}, \underline{B})$ is determined up to mixed equivalence and the other summands are uniquely determined up to permutation. This sum is called a regularizing decomposition *of (A, B). The following algorithm admits to construct a regularizing decomposition using only unitary transformations.*

Algorithm 3.3. *Let (A, B) be a pair of matrices of the same size in which the rows of A are linearly dependent. By unitary transformations of rows, we reduce A to the form*

$$S_1 A = \begin{bmatrix} 0 \\ A' \end{bmatrix}, \quad S_1 \text{ is unitary,}$$

in which the rows of A' are linearly independent. These transformations change B:

$$S_1 B = \begin{bmatrix} B' \\ B'' \end{bmatrix}.$$

By unitary transformations of columns, we reduce B' to the form $[B_1'\ 0]$ in which the columns of B_1' are linearly independent, and obtain

$$BR_1 = \begin{bmatrix} B_1' & 0 \\ \star & B_1 \end{bmatrix}, \quad R_1 \text{ is unitary.}$$

These transformations change A:

$$S_1 A \bar{R}_1 = \begin{bmatrix} 0_{k_1 l_1} & 0 \\ \star & A_1 \end{bmatrix}.$$

We apply the same procedure to (A_1, B_1) and obtain

$$(S_2 A_1 \bar{R}_2, S_2 B_1 R_2) = \left(\begin{bmatrix} 0_{k_2 l_2} & 0 \\ \star & A_2 \end{bmatrix}, \begin{bmatrix} B_2' & 0 \\ \star & B_2 \end{bmatrix} \right),$$

in which the rows of $[\star\ A_2]$ are linearly independent, S_2 and R_2 are unitary, and the columns of B_2' are linearly independent.

We repeat this procedure until we obtain

$$(S_t A_{t-1} \bar{R}_t, S_t B_{t-1} R_t) = \left(\begin{bmatrix} 0_{k_t l_t} & 0 \\ \star & A_t \end{bmatrix}, \begin{bmatrix} B_t' & 0 \\ \star & B_t \end{bmatrix} \right),$$

in which the rows of A_t are are linearly independent. The result of the algorithm is the sequence

$$(k_1, l_1),\ (k_2, l_2),\ \dots,\ (k_t, l_t),\ (A_t, B_t).$$

For a matrix pair (A, B) and a nonnegative integer n, we write

$$(A, B)^{(n)} := \begin{cases} (0_{00}, 0_{00}), & \text{if } n = 0; \\ \underbrace{(A, B) \oplus \cdots \oplus (A, B)}_{n \text{ summands}}, & \text{if } n \geqslant 1. \end{cases}$$

Theorem 3.4. *Let (A, B) be a pair of complex matrices of the same size. Let us apply Algorithm 2 to (A, B) and obtain*

$$(k_1, l_1),\ (k_2, l_2),\ \dots,\ (k_t, l_t),\ (A_t, B_t).$$

Let us apply Algorithm 2 to $(\underline{A},\underline{B}) := (B_t^T, A_t^T)$ *and obtain*

$$(\underline{k}_1,\underline{l}_1),\ (\underline{k}_2,\underline{l}_2),\ \ldots,\ (\underline{k}_{\underline{t}},\underline{l}_{\underline{t}}),\ (\underline{A}_{\underline{t}},\underline{B}_{\underline{t}}).$$

Then (A,B) *is mixed equivalent to*

$$\begin{aligned}
&(F_1,G_1)^{(k_1-l_1)}\oplus\cdots\oplus(F_{t-1},G_{t-1})^{(k_{t-1}-l_{t-1})}\oplus(F_t,G_t)^{(k_t-l_t)}\\
\oplus&(J_1,I_1)^{(l_1-k_2)}\oplus\cdots\oplus(J_{t-1},I_{t-1})^{(l_{t-1}-k_t)}\oplus(J_t,I_t)^{(l_t)}\\
\oplus&(F_1^T,G_1^T)^{(\underline{k}_1-\underline{l}_1)}\oplus\cdots\oplus(F_{\underline{t}-1}^T,G_{\underline{t}-1}^T)^{(\underline{k}_{\underline{t}-1}-\underline{l}_{\underline{t}-1})}\oplus(F_{\underline{t}}^T,G_{\underline{t}}^T)^{(\underline{k}_{\underline{t}}-\underline{l}_{\underline{t}})}\\
\oplus&(I_1,J_1)^{(\underline{l}_1-\underline{k}_2)}\oplus\cdots\oplus(I_{\underline{t}-1},J_{\underline{t}-1})^{(\underline{l}_{\underline{t}-1}-\underline{k}_{\underline{t}})}\oplus(I_{\underline{t}},J_{\underline{t}})^{(\underline{l}_{\underline{t}})}\\
\oplus&(\underline{B}_{\underline{t}}^T,\underline{A}_{\underline{t}}^T)
\end{aligned}$$

(all exponents in parentheses are nonnegative). The pair $(\underline{B}_{\underline{t}}^T,\underline{A}_{\underline{t}}^T)$ *consists of nonsingular matrices; it is determined up to mixed equivalence. The other summands are uniquely determined by* (A,B)*.*

The rows of A_t *in Theorem 3.4 are linearly independent, and so the columns of* $\underline{B} := A_t^T$ *are linearly independent. As follows from Algorithm 2, the columns of* $\underline{B}_{\underline{t}}$ *are linearly independent too. Since the rows of* $\underline{A}_{\underline{t}}$ *are linearly independent, the columns of* $\underline{B}_{\underline{t}}$ *are linearly independent, and the matrices in* $(\underline{A}_{\underline{t}},\underline{B}_{\underline{t}})$ *have the same size, these matrices are square, and so they are nonsingular. The pairs* (I_n,J_n^T) *and* (G_n^T,F_n^T) *are permutationally equivalent to* (I_n,J_n) *and* (F_n^T,G_n^T)*. Therefore, Theorem 3.4 follows from the following lemma.*

Lemma 3.1. *Let* (A,B) *be a pair of complex matrices of the same size. Let us apply Algorithm 2 to* (A,B) *and obtain*

$$(k_1,l_1),\ (k_2,l_2),\ \ldots,\ (k_t,l_t),\ (A_t,B_t).$$

Then (A,B) *is mixed equivalent to*

$$\begin{aligned}
&(F_1,G_1)^{(k_1-l_1)}\oplus\cdots\oplus(F_{t-1},G_{t-1})^{(k_{t-1}-l_{t-1})}\oplus(F_t,G_t)^{(k_t-l_t)}\\
\oplus&(J_1,I_1)^{(l_1-k_2)}\oplus\cdots\oplus(J_{t-1},I_{t-1})^{(l_{t-1}-k_t)}
\end{aligned}\qquad(3.6)$$

$$\oplus(J_t, I_t)^{(l_t)} \oplus (A_t, B_t)$$

(all exponents in parentheses are nonnegative). The rows of A_t are linearly independent. The pair (A_t, B_t) is determined up to mixed equivalence. The other summands are uniquely determined by (A, B).

Proof. We write

$$(A, B) \Longrightarrow (k_1, l_1, (A_1, B_1))$$

if $k_1, l_1, (A_1, B_1)$ are obtained from (A, B) in the first step of Algorithm 2.

First we prove two statements.

Statement 1: If

$$\begin{aligned} (A, B) &\Longrightarrow (k_1, l_1, (A_1, B_1)), \\ (\widetilde{A}, \widetilde{B}) &\Longrightarrow (\tilde{k}_1, \tilde{l}_1, (\widetilde{A}_1, \widetilde{B}_1)), \end{aligned} \tag{3.7}$$

and (A, B) is mixed equivalent to $(\widetilde{A}, \widetilde{B})$, then $k_1 = \tilde{k}_1$, $l_1 = \tilde{l}_1$, and (A_1, B_1) is mixed equivalent to $(\widetilde{A}_1, \widetilde{B}_1)$.

Let m be the number of rows in A. Then

$$k_1 = m - \operatorname{rank} A = m - \operatorname{rank} \widetilde{A} = \tilde{k}_1.$$

Since (A, B) and $(\widetilde{A}, \widetilde{B})$ are mixed equivalent and they are reduced by mixed equivalence transformations to

$$\left(\begin{bmatrix} 0_{k_1 l_1} & 0 \\ X & A_1 \end{bmatrix}, \begin{bmatrix} B_1' & 0 \\ Y & B_1 \end{bmatrix} \right), \left(\begin{bmatrix} 0_{k_1 \tilde{l}_1} & 0 \\ \widetilde{X} & \widetilde{A}_1 \end{bmatrix}, \begin{bmatrix} \widetilde{B}_1' & 0 \\ \widetilde{Y} & \widetilde{B}_1 \end{bmatrix} \right), \tag{3.8}$$

there exist nonsingular S and R such that

$$\left(S \begin{bmatrix} 0_{k_1 l_1} & 0 \\ X & A_1 \end{bmatrix}, \; S \begin{bmatrix} B_1' & 0 \\ Y & B_1 \end{bmatrix} \right) = \left(\begin{bmatrix} 0_{k_1 \tilde{l}_1} & 0 \\ \widetilde{X} & \widetilde{A}_1 \end{bmatrix} R, \; \begin{bmatrix} \widetilde{B}_1' & 0 \\ \widetilde{Y} & \widetilde{B}_1 \end{bmatrix} \bar{R} \right). \tag{3.9}$$

Equating the first matrices of these pairs, we find that S has the form

$$S = \begin{bmatrix} S_{11} & 0 \\ S_{21} & S_{22} \end{bmatrix}, \quad S_{11} \text{ is } k_1 \times k_1.$$

Equating the second matrices of the pairs (3.9), we find that

$$S_{11}[B_1' \; 0] = [\widetilde{B}_1' \; 0]\bar{R}, \tag{3.10}$$

and so

$$l_1 = \operatorname{rank}[B_1' \; 0] = \operatorname{rank}[\widetilde{B}_1' \; 0] = \tilde{l}_1.$$

Since B_1' and $\widetilde{B}_1'$ are $k_1 \times l_1$ and have linearly independent columns, (3.10) implies that R is of the form

$$R = \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix}, \quad R_{11} \text{ is } l_1 \times l_1.$$

Equating the (2,2) entries in the matrices (3.9), we get

$$S_{22}A_1 = \widetilde{A}_1 R_{22}, \qquad S_{22}B_1 = \widetilde{B}_1 \bar{R}_{22},$$

hence (A_1, B_1) and $(\widetilde{A}_1, \widetilde{B}_1)$ are mixed equivalent, which completes the proof of Statement 1.

Statement 2: If (3.7), *then*

$$(A, B) \oplus (\widetilde{A}, \widetilde{B}) \Longrightarrow (k_1 + \tilde{k}_1, l_1 + \tilde{l}_1, (A_1 \oplus \widetilde{A}_1, B_1 \oplus \widetilde{B}_1)).$$

Indeed, if (A, B) and $(\widetilde{A}, \widetilde{B})$ are reduced to (3.8), then $(A, B) \oplus (\widetilde{A}, \widetilde{B})$ is reduced to

$$\left(\begin{bmatrix} 0_{k_1 l_1} \oplus 0_{\tilde{k}_1 \tilde{l}_1} & 0 \oplus 0 \\ X \oplus \widetilde{X} & A_1 \oplus \widetilde{A}_1 \end{bmatrix}, \begin{bmatrix} B_1' \oplus \widetilde{B}_1' & 0 \oplus 0 \\ Y \oplus \widetilde{Y} & B_1 \oplus \widetilde{B}_1 \end{bmatrix} \right),$$

which is permutationally equivalent to

$$\left(\begin{bmatrix} 0_{k_1 l_1} & 0 \\ X & A_1 \end{bmatrix} \oplus \begin{bmatrix} B_1' & 0 \\ Y & B_1 \end{bmatrix} \right), \left(\begin{bmatrix} 0_{\tilde{k}_1 \tilde{l}_1} & 0 \\ \widetilde{X} & \widetilde{A}_1 \end{bmatrix} \oplus \begin{bmatrix} \widetilde{B_1'} & 0 \\ \widetilde{Y} & \widetilde{B}_1 \end{bmatrix} \right).$$

We are ready to prove Lemma 3.1 for any pair (A, B). Due to Statement 1, we can replace (A, B) by any mixed equivalent pair. In particular, we can take

$$(A, B) = (F_1, G_1)^{(r_1)} \oplus \cdots \oplus (F_t, G_t)^{(r_t)} \oplus \tag{3.11}$$

$$(J_1, I_1)^{(s_1)} \oplus \cdots \oplus (J_t, I_t)^{(s_t)} \oplus (C, D)$$

for some nonnegative $t, r_1, \ldots, r_t, s_1, \ldots, r_t$ and some pair (C, D) in which C has linearly independent rows.

Clearly,

$$(J_i, I_i) \Longrightarrow \begin{cases} (1, 1, (J_{i-1}, I_{i-1})), & \text{if } i \neq 1; \\ (1, 1, (0_{00}, 0_{00})), & \text{if } i = 1, \end{cases}$$

and

$$(F_i, G_i) \Longrightarrow \begin{cases} (1, 1, (F_{i-1}, G_{i-1})), & \text{if } i \neq 1; \\ (1, 0, (0_{00}, 0_{00})), & \text{if } i = 1. \end{cases}$$

Due to Statement 2,

- $k_1 = m - \operatorname{rank} A$ is the number of all summands of the types (J_i, I_i) and (F_i, G_i),
- l_1 is the number of all summands of the types (J_i, I_i) and (F_i, G_i), except for (F_1, G_1),
- and

$$\begin{aligned} (A_1, B_1) =& (F_1, G_1)^{(r_2)} \oplus \cdots \oplus (F_{t-1}, G_{t-1})^{(r_t)} \oplus \\ & (J_1, I_1)^{(s_2)} \oplus \cdots \oplus (J_{t-1}, I_{t-1})^{(s_t)} \oplus (C, D). \end{aligned} \tag{3.12}$$

We find that $k_1 - l_1$ is the number of summands of the type (F_1, G_1).

Applying the same reasoning to (3.12) instead of (3.11) we get that

- k_2 is the number of all summands of the types (J_i, I_i) and (F_i, G_i) with $i \geqslant 2$,
- l_1 is the number of all summands of the types (J_i, I_i) with $i \geqslant 2$ and (F_i, G_i) with $i \geqslant 3$,

- and

$$\begin{aligned}(A_2, B_2) =&(F_1, G_1)^{(r_3)} \oplus \cdots \oplus (F_{t-2}, G_{t-2})^{(r_t)} \oplus \\ &(J_1, I_1)^{(s_3)} \oplus \cdots \oplus (J_{t-2}, I_{t-2})^{(s_t)} \oplus (C, D).\end{aligned}$$

We find that $k_2 - l_2$ is the number of summands of the type (F_1, G_1), and that $l_1 - k_2$ is the number of summands of the type (J_1, I_1), and so on, until we obtain (3.6).

The fact that the pair (A_t, B_t) in (3.6) is determined up to mixed equivalence and the other summands are uniquely determined by (A, B) follows from Statement 1 (or from the canonical form of a matrix pair up to mixed equivalence). This concludes the proof of Lemma 3.1 and Theorem 2.

3.2. Perturbation analysis of the matrix differential equation $\dot{x} = ABx$

The text of this section coincides with the text of the paper [32] of M. I. García-Planas and me (up to the numeration of statements, formulas and references).

We study a matrix differential equation $\dot{x} = ABx$, whose matrix is a product of an $m \times n$ complex matrix A and an $n \times m$ complex matrix B. It is equivalent to $\dot{y} = S^{-1}ARR^{-1}BSy$, in which S and R are nonsingular matrices and $x = Sy$. Thus, we can reduce (A, B) by *transformations of contragredient equivalence*

$$(A, B) \mapsto (S^{-1}AR, R^{-1}BS), \qquad S \text{ and } R \text{ are nonsingular.} \tag{3.13}$$

The canonical form of (A, B) with respect to these transformations was obtained by Dobrovol′skaya and Ponomarev [22] and, independently, by Horn and Merino [45]:

$$\begin{aligned}&\text{each pair } (A,B) \text{ is contragrediently equivalent to a direct sum,}\\ &\text{uniquely determined up to permutation of summands, of pairs}\\ &\text{of the types } (I_r, J_r(\lambda)),\ (J_r(0), I_r),\ (F_r, G_r),\ (G_r, F_r),\end{aligned} \tag{3.14}$$

in which $r = 1, 2, \ldots,$

$$J_r(\lambda) := \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix} \ (\lambda \in \mathbb{C}),$$

$$F_r := \begin{bmatrix} 1 & & 0 \\ 0 & \ddots & \\ & \ddots & 1 \\ 0 & & 0 \end{bmatrix}, \qquad G_r := \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & 0 & 1 \end{bmatrix}$$

are $r \times r$, $r \times (r-1)$, $(r-1) \times r$ matrices, and

$$(A_1, B_1) \oplus (A_2, B_2) := (A_1 \oplus A_2, B_1 \oplus B_2).$$

Note that $(F_1, G_1) = (0_{10}, 0_{10})$; we denote by 0_{mn} the zero matrix of size $m \times n$, where $m, n \in \{0, 1, 2, \dots\}$. All matrices that we consider are complex matrices. All matrix pairs that we consider are counter pairs: a matrix pair (A, B) is a *counter pair* if A and B^T have the same size.

A notion of miniversal deformation was introduced by Arnold [3, 5]. He constructed a miniversal deformation of a Jordan matrix 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. García-Planas and Sergeichuk [39] constructed a miniversal deformation of a canonical pair (3.3.1) for contragredient equivalence (3.54).

For a counter matrix pair (A, B), we consider all matrix pairs $(A + \widetilde{A}, B + \widetilde{B})$ that are sufficiently close to (A, B). The pair $(\widetilde{A}, \widetilde{B})$ is called a *perturbation* of (A, B). Each perturbation $(\widetilde{A}, \widetilde{B})$ of (A, B) defines the *induced perturbation* $A\widetilde{B} + \widetilde{A}B + \widetilde{A}\widetilde{B}$ of the matrix AB that is obtained as follows:

$$(A + \widetilde{A})(B + \widetilde{B}) = AB + A\widetilde{B} + \widetilde{A}B + \widetilde{A}\widetilde{B}.$$

Since $\widetilde{A}$ and $\widetilde{B}$ are small, their product $\widetilde{A}\widetilde{B}$ is "very small"; we ignore it and consider only *first order induced perturbations* $A\widetilde{B} + \widetilde{A}B$ of AB.

In this paper, we describe all canonical matrix pairs (A, B) of the form (3.3.1), for which the first order induced perturbations $A\widetilde{B} + \widetilde{A}B$ are nonzero for all miniversal perturbations $(\widetilde{A}, \widetilde{B}) \neq 0$ in the normal form defined in [39].

Note that $z = ABx$ can be considered as the superposition of the systems $y = Bx$ and $z = Ay$:

$$x \longrightarrow \boxed{\;B\;} \xrightarrow{\;y\;} \boxed{\;A\;} \longrightarrow z \qquad \text{implies} \qquad x \longrightarrow \boxed{\;AB\;} \longrightarrow z$$

3.2.1. Normal form of M. I. García-Planas and V. V. Sergeichuk

In this section, we recall the miniversal deformations of canonical pairs (3.3.1) for contragredient equivalence constructed by García-Planas and Sergeichuk [39].

Let

$$\begin{aligned}(A,B) =&(I,C)\oplus\bigoplus_{j=1}^{t_1}(I_{r_{1j}},J_{r_{1j}})\oplus\bigoplus_{j=1}^{t_2}(J_{r_{2j}},I_{r_{2j}})\oplus\\ &\bigoplus_{j=1}^{t_3}(F_{r_{3j}},G_{r_{3j}})\oplus\bigoplus_{j=1}^{t_4}(G_{r_{4j}},F_{r_{4j}})\end{aligned} \tag{3.15}$$

be a canonical pair for contragredient equivalence, in which

$$C:=\bigoplus_{i=1}^{t}\Phi(\lambda_i),\qquad \Phi(\lambda_i):=J_{m_{i1}}(\lambda_i)\oplus\cdots\oplus J_{m_{ik_i}}(\lambda_i)\qquad \text{with } \lambda_i\neq\lambda_j \text{ if } i\neq j,$$

$m_{i1}\leqslant m_{i2}\leqslant\ldots\leqslant m_{ik_i}$, and $r_{i1}\leqslant r_{i2}\leqslant\ldots\leqslant r_{it_i}$.

For each matrix pair (A,B) of the form (3.29), we define the matrix pair

$$\begin{aligned}&\Big(I,\bigoplus_i(\Phi(\lambda_i)+N)\Big)\oplus\\ &\oplus\left(\left[\begin{array}{c|c|c} \oplus_j I_{r_{1j}} & 0 & 0\\ \hline 0 & \oplus_j J_{r_{2j}}(0)+N & N\\ \hline 0 & N & \begin{matrix}P_3 & N\\ 0 & Q_4\end{matrix}\end{array}\right],\left[\begin{array}{c|c|c} \oplus_j J_{r_{1j}}(0)+N & N & N\\ \hline N & \oplus_j I_{r_{2j}} & 0\\ \hline N & 0 & \begin{matrix}Q_3 & 0\\ N & P_4\end{matrix}\end{array}\right]\right),\end{aligned} \tag{3.16}$$

of the same size and of the same partition of the blocks, in which

$$N:=[H_{ij}] \tag{3.17}$$

is a parameter block matrix with $p_i\times q_j$ blocks H_{ij} of the form

$$H_{ij}:=\begin{bmatrix} * & \\ \vdots & 0\\ * & \end{bmatrix} \text{ if } p_i\leqslant q_j,\qquad H_{ij:}=\begin{bmatrix} 0\\ *\cdots*\end{bmatrix} \text{ if } p_i>q_j \tag{3.18}$$

(we usually write H_{ij} without indexes),

$$P_l := \begin{bmatrix} F_{r_{l1}} + H & H & \cdots & H \\ & F_{r_{l2}} + H & \ddots & \vdots \\ & & \ddots & H \\ 0 & & & F_{r_{lt_l}} + H \end{bmatrix},$$

$$Q_l := \begin{bmatrix} G_{r_{l1}} & & & 0 \\ H & G_{r_{l2}} & & \\ \vdots & \ddots & \ddots & \\ H & \cdots & H & G_{r_{lt_l}} \end{bmatrix} \qquad (l = 3, 4), \tag{3.19}$$

N and H are matrices of the form (3.31) and (3.32), and the stars denote independent parameters.

Theorem 3.5 (see [39]). *Let* (A, B) *be the canonical pair* (3.29). *Then all matrix pairs* $(A + \widetilde{A}, B + \widetilde{B})$ *that are sufficiently close to* (A, B) *are simultaneously reduced by some transformation*

$$(A + \widetilde{A}, B + \widetilde{B}) \mapsto (S^{-1}(A + \widetilde{A})R, R^{-1}(B + \widetilde{B})S),$$

in which S *and* R *are matrix functions that depend holomorphically on the entries of* $\widetilde{A}$ *and* $\widetilde{B}$, $S(0) = I$, *and* $R(0) = I$, *to the form* (3.30), *whose stars are replaced by complex numbers that depend holomorphically on the entries of* $\widetilde{A}$ *and* $\widetilde{B}$. *The number of stars is minimal that can be achieved by such transformations.*

3.2.2. Criteria of non-singularity: proof of the main result

Each matrix pair $(A+\widetilde{A}, B+\widetilde{B})$ of the form (3.30), in which the stars are complex numbers, we call a *miniversal normal pair* and $(\widetilde{A}, \widetilde{B})$ a *miniversal perturbation* of (A, B).

The following theorem is the main result of the paper.

Theorem 3.6. *Let (A,B) be a canonical pair* (3.3.1). *Then $A\widetilde{B}+\widetilde{A}B\neq 0$ for all nonzero miniversal perturbations $(\tilde{A},\tilde{B})$ if and only if the following inequalities hold:*

$$\begin{aligned} &r_{1t_1}<r_{21} \ \textit{ if } t_1t_2\neq 0,\\ &r_{2t_2}<r_{41} \ \textit{ if } t_2t_4\neq 0,\\ &r_{1t_1}<r_{41} \ \textit{ if } t_1t_4\neq 0, \quad \textit{and}\\ &r_{3t_3}<r_{41} \ \textit{ if } t_3t_4\neq 0.\end{aligned} \tag{3.20}$$

Proof. We write $J_r := J_r(0)$. Since the deformation (3.30) is the direct sum of

$$\Big(I,\bigoplus_i(\Phi(\lambda_i)+N)\Big)$$

and

$$\left(\left[\begin{array}{c|c|cc} \oplus_j I_{r_{1j}} & 0 & \multicolumn{2}{c}{0} \\ \hline 0 & \oplus_j J_{r_{2j}}+N & \multicolumn{2}{c}{N} \\ \hline 0 & N & P_3 & N \\ & & 0 & Q_4 \end{array}\right],\left[\begin{array}{c|c|cc} \oplus_j J_{r_{1j}}+N & N & \multicolumn{2}{c}{N} \\ \hline N & \oplus_j I_{r_{2j}} & \multicolumn{2}{c}{0} \\ \hline N & 0 & Q_3 & 0 \\ & & N & P_4 \end{array}\right]\right),$$

it is sufficient to consider (A,B) equals

$$\Big(I,\bigoplus_i\Phi(\lambda_i)\Big)$$

or

$$\bigoplus_{j=1}^{t_1}(I_{r_{1j}},J_{r_{1j}})\oplus\bigoplus_{j=1}^{t_2}(J_{r_{2j}},I_{r_{2j}})\oplus\bigoplus_{j=1}^{t_3}(F_{r_{3j}},G_{r_{3j}})\oplus\bigoplus_{j=1}^{t_4}(G_{r_{4j}},F_{r_{4j}}). \tag{3.21}$$

Let first $(A,B)=\big(I,\oplus_i\Phi(\lambda_i)\big)$. Then $(A+\widetilde{A},B+\widetilde{B})=$

$$\left(\left[\begin{array}{c|c|c} \oplus_j I_{r_{1j}} & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & \oplus_j I_{r_{lj}} \end{array}\right],\left[\begin{array}{c|c|c} \oplus_j J_{r_{1j}}(\lambda_1)+N & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & \oplus_j J_{r_{lj}}(\lambda_l)+N \end{array}\right]\right).$$

Since

$$\widetilde{A}B + A\widetilde{B} = A\widetilde{B} = \left[\begin{array}{c|c|c} N & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & N \end{array}\right],$$

in which all N have independent parameters, we have that $\widetilde{A}B + A\widetilde{B} = 0$ if and only if all N are zero, that is $(\widetilde{A}, \widetilde{B}) = (0, 0)$.

It remains to consider (A, B) equaling the second pair in (3.21). Write the matrices (3.33) as follows:

$$P_l = \overline{P}_l + \underline{P}_l, \qquad Q_l = \overline{Q}_l + \underline{Q}_l, \qquad \text{in which } \ l = 3,\, 4,$$

$$\overline{P}_l = \begin{bmatrix} F_{r_{l1}} & 0 & \cdots & 0 \\ & F_{r_{l2}} & \ddots & \vdots \\ & & \ddots & 0 \\ 0 & & & F_{r_{lt_l}} \end{bmatrix}, \qquad \underline{P}_l = \begin{bmatrix} H_{r_{l1}} & H & \cdots & H \\ & H_{r_{l2}} & \ddots & \vdots \\ & & \ddots & H \\ 0 & & & H_{r_{lt_l}} \end{bmatrix},$$

$$\overline{Q}_l = \begin{bmatrix} G_{r_{l1}} & & & 0 \\ 0 & G_{r_{l2}} & & \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & G_{r_{lt_l}} \end{bmatrix}, \qquad \underline{Q}_l = \begin{bmatrix} 0_{r_{l1}} & & & 0 \\ H & 0_{r_{l2}} & & \\ \vdots & \ddots & \ddots & \\ H & \cdots & H & 0_{r_{lt_l}} \end{bmatrix},$$

N and H are matrices of the form (3.31) and (3.32), and the stars denote independent parameters.

Write

$$\Psi_1 := \oplus_j J_{r_{1j}}(0), \qquad \Psi_2 := \oplus_j J_{r_{2j}}(0). \tag{3.22}$$

Then

$$A = \left[\begin{array}{c|c|cc} I & 0 & 0 & 0 \\ \hline 0 & \Psi_2 & 0 & 0 \\ \hline 0 & 0 & \overline{P}_3 & 0 \\ 0 & 0 & 0 & \overline{Q}_4 \end{array}\right], \qquad \widetilde{A} = \left[\begin{array}{c|c|cc} 0 & 0 & 0 & 0 \\ \hline 0 & N_{22} & N_{23} & N_{24} \\ \hline 0 & N_{32} & \underline{P}_3 & N_{34} \\ 0 & N_{42} & 0 & \underline{Q}_4 \end{array}\right],$$

$$B = \left[\begin{array}{c|c|cc} \Psi_1 & 0 & 0 & 0 \\ \hline 0 & I & 0 & 0 \\ \hline 0 & 0 & \overline{Q}_3 & 0 \\ 0 & 0 & 0 & \overline{P}_4 \end{array}\right], \qquad \widetilde{B} = \left[\begin{array}{c|c|cc} N'_{11} & N'_{12} & N'_{13} & N'_{14} \\ \hline N'_{21} & 0 & 0 & 0 \\ \hline N'_{31} & 0 & \underline{Q}_3 & 0 \\ N'_{41} & 0 & N'_{43} & \underline{P}_4 \end{array}\right],$$

$$A\widetilde{B} = \left[\begin{array}{c|c|cc} N'_{11} & N'_{12} & N'_{13} & N'_{14} \\ \hline \Psi_2 N'_{21} & 0 & 0 & 0 \\ \hline \overline{P}_3 N'_{31} & 0 & \overline{P}_3 \underline{Q}_3 & 0 \\ \overline{Q}_4 N'_{41} & 0 & \overline{Q}_4 N'_{43} & \overline{Q}_4 \underline{P}_4 \end{array}\right],$$

$$\widetilde{A}B = \left[\begin{array}{c|c|cc} 0 & 0 & 0 & 0 \\ \hline 0 & N_{22} & N_{23}\overline{Q}_3 & N_{24}\overline{P}_4 \\ \hline 0 & N_{32} & \underline{P}_3\overline{Q}_3 & N_{34}\overline{P}_4 \\ 0 & N_{42} & 0 & \underline{Q}_4\overline{P}_4 \end{array}\right],$$

in which N_{ij} and N'_{ij} are blocks of the form (3.31). All these blocks have distinct sets of independent parameters and may have distinct sizes.

Since $\widetilde{A}B$ and $A\widetilde{B}$ have independent parameters for each (A, B), we should prove that $\widetilde{A}B \neq 0$ for all $\widetilde{A} \neq 0$ and $\widetilde{B}A \neq 0$ for all $\widetilde{B} \neq 0$. Thus, we should prove that

$$\Psi_2 N'_{21}, \quad N_{23}\overline{Q}_3, \quad N_{24}\overline{P}_4, \quad \overline{P}_3 N'_{31}, \quad N_{34}\overline{P}_4, \quad \overline{Q}_4 N'_{41}, \quad \overline{Q}_4 N'_{43} \tag{3.23}$$

are nonzero if the corresponding parameter blocks N_{ij} and N'_{ij} are nonzero.

Case 1: consider the matrix

$$\Psi_2 N'_{21} = \begin{bmatrix} J_{r_{21}}(0) & & 0 \\ & \ddots & \\ 0 & & J_{r_{2t_2}}(0) \end{bmatrix} \begin{bmatrix} H_{r_{21}r_{11}} & \dots & H_{r_{21}r_{1t_1}} \\ \dots & \dots & \dots \\ H_{r_{2t_2}r_{11}} & \dots & H_{r_{2t_2}r_{1t_1}} \end{bmatrix} =$$
$$\begin{bmatrix} J_{r_{21}}(0) H_{r_{21}r_{11}} & \dots & J_{r_{21}}(0) H_{r_{21}r_{1t_1}} \\ \dots & \dots & \dots \\ J_{r_{2t_2}}(0) H_{r_{2t_2}r_{11}} & \dots & J_{r_{2t_2}}(0) H_{r_{2t_2}r_{1t_1}} \end{bmatrix}$$

in which $r_{11} \leqslant r_{12} \leqslant \cdots \leqslant r_{1t_1}$ and $r_{21} \leqslant r_{22} \leqslant \cdots \leqslant r_{2t_2}$.

The matrix N'_{21} is contained in the following submatrix of $A\widetilde{B}$:

$$\left[\begin{array}{ccc|ccc} J_{r_{11}}(0) & & 0 & & & \\ & \ddots & & & 0 & \\ 0 & & J_{r_{1t_1}}(0) & & & \\ \hline H_{r_{21}r_{11}} & \dots & H_{r_{21}r_{1t_1}} & I_{r_{21}} & & 0 \\ \vdots & & \vdots & & \ddots & \\ H_{r_{2t_2}r_{11}} & \dots & H_{r_{2t_2}r_{1t_1}} & 0 & & I_{r_{2t_2}} \end{array}\right].$$

Each $H_{r_{2i}r_{1j}}$ has the form

$$\begin{bmatrix} \alpha_{r_{21}} & \\ \vdots & 0 \\ \alpha_{r_{2i}} & \end{bmatrix} \text{ if } r_{2i} \leqslant r_{1j}, \qquad \begin{bmatrix} & 0 & \\ & & \\ \alpha_{r_{11}} & \cdots & \alpha_{r_{1j}} \end{bmatrix} \text{ if } r_{2i} > r_{1j}.$$

Correspondingly, $J_{r_{1j}}H_{r_{2i}r_{1j}}$ is

$$\begin{bmatrix} \alpha_{r_{22}} & \\ \vdots & 0 \\ \alpha_{r_{2i}-1} & \\ 0 & \end{bmatrix} \text{ if } r_{2i} \leqslant r_{1j}, \qquad \begin{bmatrix} & 0 & \\ & & \\ \alpha_{r_{11}} & \cdots & \alpha_{r_{1j}} \\ 0 & \cdots & 0 \end{bmatrix} \text{ if } r_{2i} > r_{1j}.$$

We see that $\alpha_{r_{21}}$ disappears if $r_{2i} \leqslant r_{1j}$ and all parameters remain if $r_{2i} > r_{1j}$, thus we get the inequalities $r_{11} \leqslant \cdots \leqslant r_{1t_1} < r_{21} \leqslant \cdots \leqslant r_{2t_2}$, which gives the first inequality in (3.20).

Case 2: consider the matrix

$$N_{24}\overline{P}_4 = \begin{bmatrix} H_{r_{21}r_{41}} & \dots & H_{r_{21}r_{4t_4}} \\ \dots & \dots & \dots \\ H_{r_{2t_2}r_{41}} & \dots & H_{r_{2t_2}r_{4t_4}} \end{bmatrix} \begin{bmatrix} F_{r_{41}} & & 0 \\ & \ddots & \\ 0 & & F_{r_{4t_4}} \end{bmatrix} =$$
$$\begin{bmatrix} H_{r_{21}r_{41}}F_{r_{41}} & \dots & H_{r_{21}r_{4t_4}}F_{r_{4t_4}} \\ \dots & \dots & \dots \\ H_{r_{2t_2}r_{41}}F_{r_{41}} & \dots & H_{r_{2t_2}r_{4t_4}}F_{r_{4t_4}} \end{bmatrix}$$

in which $r_{21} \leqslant \cdots \leqslant r_{2t_2}$ and $r_{41} \leqslant \cdots \leqslant r_{4t_4}$.

The matrix N_{24} is contained in the following submatrix of $\widetilde{A}B$:

$$\left[\begin{array}{ccc|ccc} J_{r_{21}}(0) & & 0 & H_{r_{21}r_{41}} & \dots & H_{r_{21}r_{4t_4}} \\ & \ddots & & \vdots & & \vdots \\ 0 & & J_{r_{2t_2}}(0) & H_{r_{2t_2}r_{41}} & \dots & H_{r_{2t_2}r_{4t_4}} \\ \hline & & & G_{r_{41}} & & 0 \\ & 0 & & & \ddots & \\ & & & 0 & & G_{r_{4t_4}} \end{array}\right].$$

Each $H_{r_{2i}r_{4j}}F_{r_{4j}}$ has the form

$$\begin{bmatrix} & & \\ & 0 & \\ \alpha_{r_{41}} & \cdots & \alpha_{r_{4j}-1} \end{bmatrix} \text{if } r_{4j} \leqslant r_{2i}, \quad \begin{bmatrix} \alpha_{r_{21}} & \\ \vdots & 0 \\ \alpha_{r_{2i}} & \end{bmatrix} \text{if } r_{4j} > r_{2i}.$$

We see that $\alpha_{r_{4j}}$ disappears if $r_{4j} \leqslant r_{2i}$ and all parameters remain if $r_{4j} > r_{2i}$, thus we have the inequalities $r_{21} \leqslant \cdots \leqslant r_{2t_2} < r_{41} \leqslant \cdots \leqslant r_{4t_4}$, which gives the second inequality in (3.20).

Case 3: consider $\overline{Q}_4 N'_{41}$. By analogy with Case 2, we get the inequalities $r_{11} \leqslant \cdots \leqslant r_{1t_1} < r_{41} \leqslant \cdots \leqslant r_{4t_4}$, which gives the third inequality in (3.20).

Case 4: consider $N_{34}\overline{P}_4$. The matrix N_{34} is contained in the following submatrix of $\widetilde{A}B$:

$$\left[\begin{array}{ccc|ccc} F_{r_{31}} & & 0 & H_{r_{31}r_{41}} & \dots & H_{r_{31}r_{4t_4}} \\ & \ddots & & \vdots & & \vdots \\ 0 & & F_{r_{3t_3}} & H_{r_{3t_3}r_{41}} & \dots & H_{r_{3t_3}r_{4t_4}} \\ \hline & & & G_{r_{41}} & & 0 \\ & 0 & & & \ddots & \\ & & & 0 & & G_{r_{4t_4}} \end{array}\right].$$

We get the inequalities $r_{31} \leqslant \cdots \leqslant r_{3t_3} < r_{41} \leqslant \cdots \leqslant r_{4t_4}$, which gives the forth inequality in (3.20).

Case 5: consider the matrix

$$N_{23}\overline{Q}_3 = \begin{bmatrix} H_{r_{21}r_{31}} & \dots & H_{r_{21}r_{3t_3}} \\ \dots & \dots & \dots \\ H_{r_{2t_2}r_{31}} & \dots & H_{r_{2t_2}r_{3t_3}} \end{bmatrix} \begin{bmatrix} G_{r_{31}} & & 0 \\ & \ddots & \\ 0 & & G_{r_{3t_3}} \end{bmatrix} =$$

$$\begin{bmatrix} H_{r_{21}r_{31}}G_{r_{31}} & \dots & H_{r_{21}r_{3t_3}}G_{r_{3t_3}} \\ \dots & \dots & \dots \\ H_{r_{2t_2}r_{31}}G_{r_{31}} & & H_{r_{2t_2}r_{3t_3}}G_{r_{3t_3}} \end{bmatrix}$$

in which $r_{21} \leqslant \dots \leqslant r_{2t_2}$ and $r_{31} \leqslant \dots \leqslant r_{3t_3}$. The matrix N_{23} is contained in the following submatrix of $\widetilde{A}B$:

$$\left[\begin{array}{ccc|ccc} J_{r_{21}}(0) & & 0 & H_{r_{21}r_{31}} & \dots & H_{r_{21}r_{3t_3}} \\ & \ddots & & \vdots & & \vdots \\ 0 & & J_{r_{2t_2}}(0) & H_{r_{2t_2}r_{31}} & \dots & H_{r_{2t_2}r_{3t_3}} \\ \hline & & & F_{r_{31}} & & 0 \\ & 0 & & & \ddots & \\ & & & 0 & & F_{r_{3t_3}} \end{array}\right].$$

Each $H_{r_{2i}r_{3j}}G_{r_{3j}}$ has the form

$$\begin{bmatrix} 0 & & \\ \vdots & & 0 \\ 0 & \alpha_{r_{21}} \cdots & \alpha_{r_{2i}} \end{bmatrix} \text{if } r_{3j} \leqslant r_{2i}, \quad \begin{bmatrix} 0 & \alpha_{r_{31}} & \\ \vdots & \vdots & 0 \\ 0 & \alpha_{r_{3j}} & \end{bmatrix} \text{if } r_{3j} > r_{2i}.$$

We find that all parameters are preserved.

Cases 6 and 7: consider the matrices $\overline{Q}_4 N'_{41}$ *and* $\overline{P}_3 N'_{31}$. We find that all parameters are preserved too.

Finally, we get that $\widetilde{A}B \neq 0$ for all $\widetilde{A} \neq 0$ and $\widetilde{B}A \neq 0$ for all $\widetilde{B} \neq 0$ if (A, B) has the form

$$\Big(I, \bigoplus_i \Phi(\lambda_i)\Big) \oplus \bigoplus_{j=1}^{t_1} (I_{r_{1j}}, J_{r_{1j}}) \oplus \bigoplus_{j=1}^{t_2} (J_{r_{2j}}, I_{r_{2j}}) \oplus \bigoplus_{j=1}^{t_3} (F_{r_{3j}}, G_{r_{3j}}) \oplus \bigoplus_{j=1}^{t_4} (G_{r_{4j}}, F_{r_{4j}})$$

in which $r_{1t_1} < r_{21}$ if $t_1 t_2 \neq 0$, $r_{2t_2} < r_{41}$ if $t_2 t_4 \neq 0$, and $r_{3t_3} < r_{41}$ if $t_3 t_4 \neq 0$.

$\square$

3.3. Versal deformations of matrix products

The text of this section coincides with the text of the paper [33] of M. I. García-Planas and me (up to the numeration of statements, formulas and references).

Let us consider the differential matrix equation $y' = ABy$, in which A and B are complex matrices. Using the substitution $y = Sz$, we can reduce (A, B) by *contragredient equivalence transformations*

$$(S^{-1}AR, R^{-1}BS), \qquad R \text{ and } S \text{ are nonsingular} \tag{3.24}$$

since $y' = ABy$ is equivalent to $z' = S^{-1}AR \cdot R^{-1}BSz$. We study perturbations of AB that are products of perturbations of A and B.

By Arnold [3], a *deformation* of a square complex matrix M is a matrix $M(\delta)$, $\delta = (\delta_1, \dots, \delta_r)$ of the same size with entries that are power series of an arbitrary number of complex variables $\delta_1, \dots, \delta_r$, convergent in a neighborhood of 0, with $M(0) = M$. A *simplest deformation* of M is a deformation $M+X$, in which every entry of X is either 0, or an independent parameter x_{ij}.

A deformation $M(\delta)$ is *versal under similarity* if all complex matrices $M + E$ that are sufficiently close to M can be simultaneously reduced by some transformation

$$S(E)^{-1}(M + E)S(E), \quad S(E) \text{ is holomorphic at zero, } S(0) = I$$

to the form $M(\delta_1(E), \dots, \delta_r(E))$, in which all $\delta_i(E)$ are holomorphic functions on the entries of E such that $\delta_i(0) = 0$. A versal deformation with the minimal number of parameters is called *miniversal*. For example, the simplest deformation $M(X) = M+X$ in which all entries of X are independent parameters is versal. Arnold [3, 5] constructed miniversal deformations of all Jordan matrices; in particular, the parameter matrix

$$J_m(\lambda)+X=\begin{bmatrix}\lambda&1&&0\\&\lambda&\ddots&\\&&\ddots&1\\0&&&\lambda\end{bmatrix}+\begin{bmatrix}0&0&\dots&0\\\vdots&\vdots&&\vdots\\0&0&\dots&0\\x_{m1}&x_{m2}&\dots&x_{mm}\end{bmatrix}\qquad(\lambda\in\mathbb{C})$$

is a miniversal simplest deformation of the Jordan block $J_m(\lambda)$.

Versal deformations provide a special parametrization of matrix spaces, which can be effectively applied to perturbation analysis and investigation of complicated objects like singularities and bifurcations in multiparameter dynamical systems; see [35, 56] and the references given there. Versal deformations are widely used both in abstract mathematics and in its applications. For example, Bleher and Chinburg [7] apply versal deformations to Galois theory; Conrad [14] applies them to algebraic geometry; Koçak [49] constructs versal deformations of linear Hamiltonian systems.

A deformation $(A(\delta),B(\delta))$ of a pair (A,B) of complex matrices of sizes $m\times n$ and $n\times m$ and its versality under contragredient equivalence (3.54) are determined analogously. García-Planas and Sergeichuk [39] constructed miniversal simplest deformations of matrix pairs under contragredient equivalence.

Let $(A(\delta),B(\delta))$ be a versal deformation of (A,B). It is important to know, when the deformation $A(\delta)B(\delta)$ of AB is versal, which means that the behavior of AB under perturbations is fully determined by the deformation $(A(\delta),B(\delta))$. In particular, if $(A+X,B+Y)$ is the simplest deformation of (A,B), in which all entries of X and Y are independent parameters, then

$$(A+X)(B+Y)=AB+AY+XB+XY \tag{3.25}$$

is a deformation of AB and

- if $m\leqslant n$, then (3.25) is a versal deformation since the matrix equation $AY+XB+XY=E$ is solvable for each E (we can take a small X

such that the rows of $A+X$ are linearly independent and find Y from $(A+X)Y = E - XB)$;

- if m is much more than n, then (3.25) is not a versal deformation (we do not know if the condition $m > n$ is sufficient).

The main result of this paper is Theorem 3.9, in which we prove the equivalence of two statements for a pair (A, B) of complex matrices of sizes $m \times n$ and $n \times m$:

- if $(A + X, B + Y)$ is the simplest miniversal deformation of (A, B) constructed by García-Planas and Sergeichuk [39], then $(A+X)(B+Y)$ is a versal deformation of AB;

- A has linearly independent rows and/or B has linearly independent columns.

3.3.1. Preliminaries For each Jordan canonical matrix J whose Jordan blocks are ordered as follows:

$$J = \bigoplus_{i=1}^{t}(J_{m_{i1}}(\lambda_i) \oplus \cdots \oplus J_{m_{ir_i}}(\lambda_i)), \qquad m_{i1} \geqslant m_{i2} \geqslant \ldots \geqslant m_{ir_i} \tag{3.26}$$

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

$$J + \mathcal{D} := \bigoplus_{i=1}^{t} \begin{bmatrix} J_{m_{i1}}(\lambda_i) + 0^{\downarrow} & 0^{\downarrow} & \ldots & 0^{\downarrow} \\ 0^{\leftarrow} & J_{m_{i2}}(\lambda_i) + 0^{\downarrow} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0^{\downarrow} \\ 0^{\leftarrow} & \ldots & 0^{\leftarrow} & J_{m_{ir_i}}(\lambda_i) + 0^{\downarrow} \end{bmatrix} \tag{3.27}$$

in which

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

and the stars denote independent parameters.

Theorem 3.7 (Arnold [3]). *The deformation* (3.27) *of the Jordan canonical matrix* (3.26) *is miniversal.*

The canonical form of a pair (A, B) of $m \times n$ and $n \times m$ complex matrices under contragredient equivalence is given by Dobrovol′skaya and Ponomarev [22] and, independently, by Horn and Merino [45]:

> (A, B) is contragrediently equivalent to a direct sum, uniquely determined up to permutation of summands, of pairs of the types $(I_r, J_r(\lambda))$, $(J_r(0), I_r)$, (F_r, G_r), (G_r, F_r),

in which $r = 1, 2, \dots,$

$$F_r := \begin{bmatrix} 1 & & 0 \\ 0 & \ddots & \\ & \ddots & 1 \\ 0 & & 0 \end{bmatrix}, \qquad G_r := \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & 0 & 1 \end{bmatrix} \tag{3.28}$$

are $r \times (r-1)$ and $(r-1) \times r$ matrices, and

$$(A_1, B_1) \oplus (A_2, B_2) := (A_1 \oplus A_2, B_1 \oplus B_2).$$

Note that

$$(F_1, G_1) = (0_{10}, 0_{01}).$$

We denote by 0_{mn} the $m \times n$ zero matrix, where $m, n \in \{0, 1, 2, \dots\}$.

Let us write a canonical pair for contragredient equivalence in the form

$$\begin{aligned}(A, B) = (I, C) \oplus \bigoplus_{j=1}^{t_1} (I_{r_{1j}}, J_{r_{1j}}) \oplus \bigoplus_{j=1}^{t_2} (J_{r_{2j}}, I_{r_{2j}}) \\ \oplus \bigoplus_{j=1}^{t_3} (F_{r_{3j}}, G_{r_{3j}}) \oplus \bigoplus_{j=1}^{t_4} (G_{r_{4j}}, F_{r_{4j}}),\end{aligned} \tag{3.29}$$

in which

$$C := \bigoplus_{i=1}^{t} \Phi(\lambda_i), \qquad \Phi(\lambda_i) := J_{m_{i1}}(\lambda_i) \oplus \cdots \oplus J_{m_{ik_i}}(\lambda_i)$$

with $\lambda_i \neq \lambda_j$ if $i \neq j$, $\quad m_{i1} \geqslant m_{i2} \geqslant \cdots \geqslant m_{ik_i}$, $\quad$ and $\quad r_{i1} \geqslant r_{i2} \geqslant \cdots \geqslant r_{it_i}$.

For each matrix pair (A, B) of the form (3.29), we define the matrix pair

$$\Big(I, \bigoplus_i (\Phi(\lambda_i)+N)\Big) \oplus \left(\left[\begin{array}{c|c|cc} \oplus_j I_{r_{1j}} & 0 & \multicolumn{2}{c}{0} \\ \hline 0 & \oplus_j J_{r_{2j}}(0)+N & \multicolumn{2}{c}{N} \\ \hline 0 & N & P_3 & N \\ & & 0 & Q_4 \end{array}\right], \right.$$
$$\left. \left[\begin{array}{c|c|cc} \oplus_j J_{r_{1j}}(0)+N & N & \multicolumn{2}{c}{N} \\ \hline N & \oplus_j I_{r_{2j}} & \multicolumn{2}{c}{0} \\ \hline N & 0 & Q_3 & 0 \\ & & N & P_4 \end{array}\right] \right), \qquad (3.30)$$

of the same size and of the same partition of the blocks, in which

$$N := [H_{ij}] \qquad (3.31)$$

is a parameter block matrix with $p_i \times q_j$ blocks H_{ij} of the form

$$H_{ij} := \begin{bmatrix} * & \\ \vdots & 0 \\ * & \end{bmatrix} \text{ if } p_i \leqslant q_j, \qquad H_{ij:} = \begin{bmatrix} 0 \\ * \cdots * \end{bmatrix} \text{ if } p_i > q_j \qquad (3.32)$$

(we usually write H_{ij} without indexes),

$$P_l := \begin{bmatrix} F_{r_{l1}}+H & H & \cdots & H \\ & F_{r_{l2}}+H & \ddots & \vdots \\ & & \ddots & H \\ 0 & & & F_{r_{lt_l}}+H \end{bmatrix}, \quad Q_l := \begin{bmatrix} G_{r_{l1}} & & & 0 \\ H & G_{r_{l2}} & & \\ \vdots & \ddots & \ddots & \\ H & \cdots & H & G_{r_{lt_l}} \end{bmatrix}, \qquad (3.33)$$

($l = 3, 4$), N and H are matrices of the form (3.31) and (3.32), and the stars denote independent parameters.

Theorem 3.8 (Garcia–Planas, Sergeichuk [39]). *The deformation* (3.30) *of the canonical pair* (3.29) *is miniversal.*

3.3.2. Main theorem Let us define the matrix pairs

$$\mathcal{J}_n(\lambda) := (I_n, J_n(\lambda)), \qquad \mathcal{K}_n := (J_n(0), I_n),$$

$$\mathcal{F}_n := (F_n, G_n), \qquad \mathcal{G}_n := (G_n, F_n),$$

in which the matrices F_n and G_n are of the form (3.28).

The following theorem is the main result of the paper.

Theorem 3.9. *Let (A, B) be the canonical pair (3.29). Then the following three conditions are equivalent:*

(a) *if $(A + X, B + Y)$ is the simplest miniversal deformation (3.30) of (A, B), then $(A + X)(B + Y)$ is a versal deformation of AB;*

(b) *(A, B) does not contain summands of types $\mathcal{F}_r$ and $\mathcal{J}_m(0) \oplus \mathcal{K}_n$.*

(c) *A has linearly independent rows and/or B has linearly independent columns.*

Lemma 3.2. *Let*

$$M = J_{m_1}(\lambda_1) \oplus \cdots \oplus J_{m_k}(\lambda_k). \tag{3.34}$$

The tangent space to the orbit of M at the point M is equal to

$$T(M) := \{SM - MS \,|\, S \in \mathbb{C}^{n\times n}\}. \tag{3.35}$$

$T(M)$ consists of all the block matrices

$$P := [P_{ij}]_{i,j=1}^{k}$$

divided into blocks conformally with M, in which

- *if $\lambda_i \neq \lambda_j$ then P_{ij} is an arbitrary matrix;*

- *if $\lambda_i = \lambda_j$ then*

$$P_{ij} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ b_1 & \ddots & & \vdots \\ \ddots & b_2 & * & * \\ a_1 & \ddots & \ddots & * \\ 0 & a_2 & \ddots & b_{m_j} \end{bmatrix} \text{ if } m_i \geqslant m_j \tag{3.36}$$

and

$$P_{ij} = \begin{bmatrix} b_1 & * & \dots & \dots & \dots & \dots & * \\ \ddots & b_2 & \ddots & & & & \vdots \\ a_1 & \ddots & \ddots & \ddots & & & \vdots \\ 0 & a_2 & \ddots & b_{m_i} & * & \dots & * \end{bmatrix} \text{ if } m_i < m_j, \tag{3.37}$$

in which the sum of elements of each first $\min(m_i, m_j)$ *diagonals starting from the left bottom corner equals zero:*

$$a_1 + a_2 = 0, \; \dots, \; b_1 + b_2 + \cdots + b_{min(m_i,m_j)} = 0,$$

and the elements of the other diagonals (denoted by $*$*) are arbitrary.*

Proof. The tangent space to the orbit of M at the point M has the form (3.35) since by the Lipschitz property [61] each matrix that is similar to M and close to M has the following form with a small S:

$$\begin{aligned}(I-S)^{-1}M(I-S) &= (I+S+S^2+\cdots)M(I-S) \\ &= M + (SM-MS) + S(SM-MS) + S^2(SM-MS) + \cdots \\ &= M + \underbrace{SM-MS}_{\text{small}} + \underbrace{S(I-S)^{-1}(SM-MS)}_{\text{very small}}.\end{aligned}$$

A matrix P belongs to $T(M)$ if and only if there exists S such that

$$SM - MS = P. \tag{3.38}$$

Equating the corresponding blocks in (3.38), we obtain the system of k^2 equalities:

$$S_{ij}J_{m_i}(\lambda_i) - J_{m_j}(\lambda_j)S_{ij} = P_{ij}, \tag{3.39}$$

in which $i, j = 1, 2, \dots, k$.

If $\lambda_i \neq \lambda_j$ in (3.34) then for each P_{ij} there exists S_{ij} such that (3.39) holds (see [30, Chapter VIII, § 3]).

If $\lambda_i = \lambda_j$ then

$$S_{ij}(\lambda_i I + J_{m_i}(0)) - (\lambda_i I + J_{m_j}(0))S_{ij} = P_{ij},$$

which gives

$$S_{ij}J_{m_i}(0) - J_{m_j}(0)S_{ij} = P_{ij}. \tag{3.40}$$

Write

$$S_{ij} = \begin{bmatrix} s_{11} & \dots & s_{1m_j} \\ \vdots & \ddots & \vdots \\ s_{m_i1} & \dots & s_{m_im_j} \end{bmatrix}.$$

Then (3.40) takes the form

$$\begin{bmatrix} 0 & s_{11} & \dots & s_{1,m_j-1} \\ 0 & s_{21} & \dots & s_{2,m_j-1} \\ \vdots & \vdots & & \vdots \\ 0 & s_{m_i1} & \dots & s_{m_i,m_j-1} \end{bmatrix} - \begin{bmatrix} s_{21} & s_{22} & \dots & s_{2m_j} \\ \vdots & \vdots & & \vdots \\ s_{m_i1} & s_{m_i2} & \dots & s_{m_im_j} \\ 0 & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} p_{11} & \dots & p_{1m_j} \\ \vdots & & \vdots \\ p_{m_i1} & \dots & p_{m_im_j} \end{bmatrix},$$

which proves that P_{ij} has the form (3.36) or (3.37). □

Recall that a deformation of a square matrix M is a power series

$$M(\delta_1, \dots, \delta_r) = M + \sum_i M_i\delta_i + \sum_{ij} M_{ij}\delta_i\delta_j + \dots.$$

that is convergent in a neighborhood of 0. Its *linearization* is the deformation

$$M(\delta_1, \dots, \delta_r) = M + \sum_i M_i\delta_i.$$

Note that Arnold's deformation (3.27) coinsides with its linearization.

Define the vector space

$$V(M(\delta_1, \dots, \delta_r)) := \Big\{ \sum_i M_i a_i \,\Big|\, a_1, \dots a_r \in \mathbb{C} \Big\}. \tag{3.41}$$

Lemma 3.3 (Arnold [3,5]). *A deformation $M(\delta_1, \dots, \delta_r)$ of an $n\times n$ matrix M is versal if and only if*

$$T(M) + V(M(\delta_1, \dots, \delta_r)) = \mathbb{C}^{n\times n}.$$

Proof. of Theorem 3.9 (a)$\Longrightarrow$(b) Let (A, B) be the matrix pair (3.29) and let $(A + X, B + Y)$ be its simplest miniversal deformation (3.30). Let us prove that if (A, B) contains the summand $\mathcal{F}_r$ or $\mathcal{J}_m(0) \oplus \mathcal{K}_n$, then the deformation $(A+X)(B+Y)$ of AB is not versal.

Multiplying the horizontal strip of $A+X$ that contains the block P_3 by the vertical strip of $B+Y$ that contains the block Q_3, we get

$$P_3 Q_3 + N_1 N_2, \tag{3.42}$$

in which P_3 and Q_3 are defined in (3.33); N_3 and N_4 are parameter block matrices of the form (3.31).

The pair (P_3, Q_3) is a versal deformation of $\bigoplus_{i=1}^{t_3} \mathcal{F}_{r_{3i}}$. Suppose that the deformation $(A+X)(B+Y)$ is versal. By Lemma 3.3, the sum of the tangent space to the orbit AB and of the vector space defined by the linearization of this deformation is equal to the vector space of all matrices of this size. Then the sum of the tangent space to the orbit of $\bigoplus_{i=1}^{t_3} F_{r_{3i}} G_{r_{3i}}$ and of the vector space defined by the linearization of (3.42) is equal to the vector space of all matrices of this size. The linearization of (3.42) coinsides with the linearization of $P_3 Q_3$. Therefore, the deformation $P_3 Q_3$ is versal.

Multiplying the horizontal strip of P_3 that contains the block $F_{r_1} + H$ by the vertical strip of Q_3 that contains the block G_{r_1}, we get

$$(F_{r_1} + H) G_{r_1} + H_{1 r_2} H_{r_2 1} + \cdots + H_{1 r_{t_3}} H_{r_{t_3} 1}. \tag{3.43}$$

Since the deformation $P_3 Q_3$ is versal, by Lemma 3.3 the sum of the tangent space to the orbit of $\bigoplus_{i=1}^{t_3} F_{r_{3i}} G_{r_{3i}}$ and of the vector space defined by the linearization of $P_3 Q_3$ is equal to the vector space of all matrices of this size. Then the sum of the tangent space to the orbit of $F_{r_1} G_{r_1}$ and of the vector space defined by the linearization of (3.43) is equal to the vector space of all matrices of this size. Therefore, the deformation (3.43) is versal. Since the linearization of (3.43) coincides with $(F_{r_1} + H) G_{r_1}$, the deformation

$(F_{r_1}+H)G_{r_1}$ is versal too. It has the form

$$(F_{r_1}+H)G_{r_1} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & \\ \vdots & & \ddots & \ddots & \\ 0 & & & 0 & 1 \\ 0 & \alpha_1 & \alpha_2 & \cdots & \alpha_{r_1-1} \end{bmatrix}. \tag{3.44}$$

in which $\alpha_1, \dots, \alpha_{r_1-1}$ are independent parameters.

By Lemma 3.2, the left bottom entry of all matrices in the tangent space to the orbit of $F_{r_1}G_{r_1} = J_{r_1}(0)$ is zero. The left bottom entry of (3.44) is zero too. Therefore the sum of the tangent space to the orbit of $F_{r_1}G_{r_1}$ and of the vector space defined by (3.44) is not equal to the vector space of all matrices of this size. Therefore, the deformation (3.44) is not versal, and so the deformation $(A+X)(B+Y)$ is not versal too, which proves the implication (a)$\Longrightarrow$(b) for pairs $\mathcal{F}_n$.

Multiplying the horizontal strip of $A+X$ that contains the block $(\oplus_j I_{r_{1j}}, (\oplus_j J_{r_{2j}}(0) + N_{r_2}))$ by the vertical strip of $B+Y$ that contains the block $((\oplus_j J_{r_{1j}}(0) + N_{r_1}), \oplus_j I_{r_{2j}})$, we get

$$\begin{bmatrix} \oplus_j J_{r_{1j}}(0) + N_{r_1} & N_1 \\ (\oplus_j J_{r_{2j}}(0) + N_{r_2})N_2 + N_3N_4 & \oplus_j J_{r_{2j}}(0) + N_{r_2} \end{bmatrix} \tag{3.45}$$

in which all N_i are parameter block matrices (3.30).

The pair

$$\left(\begin{bmatrix} \oplus_j I_{r_{1j}} & 0 \\ 0 & \oplus_j J_{r_{2j}}(0) + N_{r_2} \end{bmatrix}, \begin{bmatrix} \oplus_j J_{r_{1j}}(0) + N_{r_1} & N_1 \\ N_2 & \oplus_j I_{r_{2j}} + N_{r_2} \end{bmatrix}\right)$$

is a versal deformation of $(\oplus_j I_{r_{1j}}, \oplus_j J_{r_{1j}}(0)) \bigoplus (\oplus_j J_{r_{2j}}(0), \oplus_j I_{r_{2j}})$. Suppose that the deformation $(A+X)(B+Y)$ is versal. By Lemma 3.3, the sum of the tangent space to the orbit AB and of the vector space defined by the linearization of this deformation is equal to the vector space of all matrices of this size. Then the sum of the tangent space to the orbit of $(\oplus_j I_{r_{1j}} \bigoplus \oplus_j J_{r_{2j}}(0))(\oplus_j J_{r_{1j}}(0) \bigoplus \oplus_j I_{r_{2j}})$ and of the vector space defined by the linearization of (3.45) is equal to the vector space of all matrices of this size. The linearization of (3.45) coinsides with the linearization of $(\oplus_j I_{r_{1j}} \bigoplus(\oplus_j J_{r_{2j}} + N_{r_2}(0)))((\oplus_j J_{r_{1j}} + N_{r_1}(0)) \bigoplus \oplus_j I_{r_{2j}})$. Therefore, the deformation $(\oplus_j I_{r_{1j}} \bigoplus(\oplus_j J_{r_{2j}} + N_{r_2}(0)))((\oplus_j J_{r_{1j}} + N_{r_1}(0)) \bigoplus \oplus_j I_{r_{2j}})$ is versal.

Multiplying the horizontal strip of $\oplus_j I_{r_{1j}} \bigoplus(\oplus_j J_{r_{2j}}(0) + N_{r_2})$ that contains the block $I_{r_1} \oplus (J_{r_2}(0) + H_{r_2})$ by the vertical strip of $(\oplus_j J_{r_{1j}}(0) + N_{r_1}) \bigoplus \oplus_j I_{r_{2j}}$ that contains the block $(J_{r_1}(0) + H_{r_1}) \oplus I_{r_2}$, we get

$$\begin{bmatrix} J_{r_1}(0) + H_{r_1} & H_1 \\ (J_{r_2}(0) + H_{r_2})H_2 + H_3H_4 + \cdots + H_{r_3}H_{r_4} & J_{r_2}(0) + H_{r_2} \end{bmatrix} \tag{3.46}$$

Since the deformation $(\oplus_j I_{r_{1j}} \bigoplus(\oplus_j J_{r_{2j}} + N_{r_2}(0)))$ $((\oplus_j J_{r_{1j}} + N_{r_1}(0)) \bigoplus \oplus_j I_{r_{2j}})$ is versal, by Lemma 3.3 the sum of the tangent space to the orbit of $(\oplus_j I_{r_{1j}} \bigoplus \oplus_j J_{r_{2j}}(0))$ $(\oplus_j J_{r_{1j}}(0) \bigoplus \oplus_j I_{r_{2j}})$ and of the vector space defined by the linearization of $(\oplus_j I_{r_{1j}} \bigoplus(\oplus_j J_{r_{2j}} + N_{r_2}(0)))((\oplus_j J_{r_{1j}} + N_{r_1}(0)) \bigoplus \oplus_j I_{r_{2j}})$ is equal to the vector space of all matrices of this size. Then the sum of the tangent space to the orbit of $(I_{r_1} \oplus J_{r_2}(0))(J_{r_1}(0) \oplus I_{r_2})$ and of the vector space defined by the linearization of (3.46) is equal to the vector space of all matrices of this size. Therefore, the deformation (3.46) is versal. Since the linearization of (3.46) coincides with

$$\begin{bmatrix} J_{r_1}(0) + H_{r_1} & H_1 \\ J_{r_2}(0) & J_{r_2}(0) + H_{r_2} \end{bmatrix}, \tag{3.47}$$

the deformation (3.47) is versal too.

By Lemma 3.2, the left bottom entry of all matrices in the tangent space to the orbit of $(\oplus_j I_{r_{1j}} \bigoplus \oplus_j J_{r_{2j}}(0))$ $(\oplus_j J_{r_{1j}}(0) \bigoplus \oplus_j I_{r_{2j}})$ is zero. The left bottom entry of (3.47) is zero too. Therefore the sum of the tangent space to the orbit of $(\oplus_j I_{r_{1j}} \bigoplus \oplus_j J_{r_{2j}}(0))$ $(\oplus_j J_{r_{1j}}(0) \bigoplus \oplus_j I_{r_{2j}})$ and of the vector space defined by (3.47) is not equal to the vector space of all matrices of this size. Therefore, the deformation (3.47) is not versal, and so the deformation $(A+X)(B+Y)$ is not versal too, which proves the implication (a)$\Longrightarrow$(b) for pairs $\mathcal{K}_n$.

(b)$\Longrightarrow$(a) Let us prove that if (A,B) of the form (3.29) does not contain summands of types $\mathcal{F}_r$ and $\mathcal{J}_m(0)\oplus\mathcal{K}_n$, then it is versal.

Case 1: *(A,B) contains summands of type $\mathcal{J}_n(0)$.* Then it does not contain summands of type $\mathcal{K}_n$, and so its miniversal deformation is the direct sum of

$$
\begin{aligned}
&(A_1+X_1, B_1+Y_1)\\
&\quad=\left(\left[\begin{array}{c|c|c} \oplus_j I_{r_{1j}} & 0 & 0\\ \hline 0 & \ddots & 0\\ \hline 0 & 0 & \oplus_j I_{r_{lj}}\end{array}\right],\left[\begin{array}{c|c|c} \oplus_j J_{r_{1j}}(\lambda_1)+N & 0 & 0\\ \hline 0 & \ddots & 0\\ \hline 0 & 0 & \oplus_j J_{r_{lj}}(\lambda_l)+N\end{array}\right]\right)
\end{aligned}
$$

and

$$
(A_2+X_2, B_2+Y_2)=\left(\left[\begin{array}{cc} c|c\oplus_j I_{r_{1j}} & 0\\ \hline 0 & Q_4\end{array}\right],\left[\begin{array}{cc} c|c\oplus_j J_{r_{1j}}+N_{r_1} & N_1\\ \hline N_2 & P_4\end{array}\right]\right).
$$

It is sufficient to prove that $(A_1+X_1)(B_1+Y_1)$ and $(A_2+X_2)(B_2+Y_2)$ are versal. The deformation $(A_1+X_1)(B_1+Y_1)$ is the miniversal deformation of the Jordan matrix $AB=\oplus_j J_{r_j}(\lambda)$.

The deformation $(A_2+X_2)(B_2+Y_2)$ has the form

$$
\left[\begin{array}{cc} c|c\oplus_j J_{r_{1j}}+N_{r_1} & N_1\\ \hline Q_4N_2 & Q_4P_4\end{array}\right], \tag{3.48}
$$

in which N_i, H, P_4, and Q_4 are defined in (3.31)–(3.33).

The space (3.41) defined by the linearization of deformation (3.48) contains the space defined by the linear deformation of A_2B_2 given in (3.27). Since the deformation (3.27) is versal, then by Lemma 3.3 the deformation (3.48) is versal too.

Case 2: (A, B) *contains summands of type* $\mathcal{K}_n$. Then it does not contain summands of type $\mathcal{J}_n(0)$, and so its miniversal deformation is the direct sum of

$$(A_1 + X_1, B_1 + Y_1)$$
$$= \left(\left[\begin{array}{c|c|c} \oplus_j I_{r_{1j}} & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & \oplus_j I_{r_{lj}} \end{array} \right], \left[\begin{array}{c|c|c} \oplus_j J_{r_{1j}}(\lambda_1) + N & 0 & 0 \\ \hline 0 & \ddots & 0 \\ \hline 0 & 0 & \oplus_j J_{r_{lj}}(\lambda_l) + N \end{array} \right] \right)$$

and

$$(A_2 + X_2, B_2 + Y_2) = \left(\left[\begin{array}{cc} c|c \oplus_j J_{r_{2j}} + N_{r_2} & N_1 \\ \hline N_2 & Q_4 \end{array} \right], \left[\begin{array}{cc} c|c \oplus_j I_{r_{2j}} & 0 \\ \hline 0 & P_4 \end{array} \right] \right).$$

It is sufficient to prove that $(A_1 + X_1)(B_1 + Y_1)$ and $(A_2 + X_2)(B_2 + Y_2)$ are versal. The deformation $(A_1 + X_1)(B_1 + Y_1)$ is the miniversal deformation of the Jordan matrix $AB = \oplus_j J_{r_j}(\lambda)$.

The deformation $(A_2 + X_2)(B_2 + Y_2)$ has the form

$$\left[\begin{array}{cc} c|c \oplus_j J_{r_{2j}} + N_{r_2} & N_1 P_4 \\ \hline N_2 & Q_4 P_4 \end{array} \right] \tag{3.49}$$

in which N_i, H, P_4, and Q_4 are defined in (3.31)–(3.33).

The space (3.41) defined by the linearization of deformation (3.49) contains the space defined by the linear deformation of $A_2 B_2$ given in (3.27). Since the deformation (3.27) is versal, then by Lemma 3.3 the deformation (3.49) is versal too.

(b)$\Longleftrightarrow$(c) This equivalence is obvious. $\square$

3.4. Structural stability of matrix pairs under contragredient equivalence

The text of this section coincides with the text of the paper [34] of M. I. García-Planas and me (up to the numeration of statements, formulas and references).

Each matrix problem $\mathcal{M}$ over $\mathbb{C}$ is given by a set $\mathcal{M}_0$ of tuples of complex matrices and a set $\mathcal{M}_1$ of admissible transformations with them. A matrix tuple $\mathcal{A} = (A_1, \dots, A_t) \in \mathcal{M}_0$ is *structurally stable* if each matrix tuple $\mathcal{B} \in \mathcal{M}_0$ that is sufficiently close to $\mathcal{A}$ can be reduced to $\mathcal{A}$ by admissible transformations. This notion is used in [37, 41, 74]. It is inspired by the notion of structurally stable dynamical systems given by Andronov and Pontryagin [2] (see also [54, 62]).

For example, an $m \times n$ matrix is structurally stable under elementary transformations if and only if its rank is $\min(m, n)$. Each square matrix A is structurally unstable under similarity since its eigenvalues can be changed by an arbitrarily small perturbation.

In Section 3.4.1, we describe all pairs (A, B) of $m \times n$ complex matrices that are structurally stable with respect to equivalence transformations $(S^{-1}AR, S^{-1}BR)$ with nonsingular R and S (that is, we describe all pencils $A - \lambda B$ that are structurally stable). In Section 3.4.2, we show that there are no pairs (M, N) of $m \times n$ and $n \times m$ complex matrices that are structurally stable under contragredient equivalence $(S^{-1}MR, R^{-1}NS)$. Perturbations of matrix pairs under equivalence and contragredient equivalence are studied in [39, 40, 48].

3.4.1. Matrix pairs that are structurally stable with respect to equivalence Each matrix pair in this section consists of two complex matrices whose sizes are equal.

The problem of classifying complex matrix pencils $A-\lambda B$ is the problem of classifying pairs of complex matrices of the same size up to equivalence transformations

$$(A,B)\mapsto(S^{-1}AR,S^{-1}BR),\qquad S \text{ and } R \text{ are nonsingular.} \tag{3.50}$$

By Kronecker's theorem for matrix pencils (see [28, Section 1.8]), (A,B) is equivalent to a direct sum, uniquely determined up to permutation of summands, of pairs of the types

$$(I_r,J_r(\lambda)),\ (J_r(0),I_r),\ (L_r,R_r),\ (L_r^T,R_r^T),\quad r=1,2,\dots, \tag{3.51}$$

in which

$$J_r(\lambda):=\begin{bmatrix}\lambda&&&0\\1&\lambda&&\\&\ddots&\ddots&\\0&&1&\lambda\end{bmatrix}\quad (r\text{-by-}r,\ \lambda\in\mathbb{C}),$$

$$L_r:=\begin{bmatrix}1&0&&0\\&\ddots&\ddots&\\0&&1&0\end{bmatrix},\quad R_r:=\begin{bmatrix}0&1&&0\\&\ddots&\ddots&\\0&&0&1\end{bmatrix}\quad ((r-1)\text{-by-}r). \tag{3.52}$$

Note that $(L_1,R_1)=(0_{10},0_{01})$; we denote by 0_{mn} the zero matrix of size $m\times n$, in which $m,n=0,1,2,\dots$. The direct sum of matrix pairs is defined by

$$(A,B)\oplus(A',B'):=(A\oplus A',B\oplus B').$$

Theorem 3.10. *A pair (A,B) of complex matrices of the same size is structurally stable under equivalence if and only if (A,B) or (A^T,B^T) is equivalent to a pair of the form*

$$\underbrace{(L_r,R_r)\oplus\dots\oplus(L_r,R_r)}_{p\ times}\oplus\underbrace{(L_{r+1},R_{r+1})\oplus\dots\oplus(L_{r+1},R_{r+1})}_{q\ times}, \tag{3.53}$$

in which $p\geqslant 1$ and $q\geqslant 0$.

Proof. $\Longrightarrow$. Let (A,B) be a direct sum of pairs of the form (3.51), and let (A,B) be structurally stable. Clearly, it does not contain direct summands of the form $(I_r, J_r(\lambda))$ and $(J_r(0), I_r)$. It does not contain $(L_r, R_r) \oplus (L_s^T, R_s^T)$ since $L_r \oplus L_s$ is a square singular matrix, which can be made nonsingular by an arbitrarily small perturbation. Therefore, (A,B) is a direct sum of summands of the form (L_r, R_r), or a direct sum of summands of the form (L_r^T, R_r^T).

Let (A,B) be a direct sum of summands of the form (L_r, R_r), and let (A,B) be not of the form (3.53). Then (A,B) has a direct summand $(L_r, R_r) \oplus (L_s, R_s)$ with $s - r \geqslant 2$. Pokrzywa [60, Theorem 3] describes inclusion relationships between the closures of the equivalence classes of two matrix pencils. Using his theorem in the form presented in [19, Theorem 2.2], we find that $(L_r, R_r) \oplus (L_s, R_s)$ is reduced to a pair that is equivalent to $(L_{r+1}, R_{r+1}) \oplus (L_{s-1}, R_{s-1})$ by an arbitrarily small perturbation of $(L_r, R_r) \oplus (L_s, R_s)$, which contradicts to the structural stability of (A,B). Hence, (A,B) is of the form (3.53).

$\Longleftarrow$. Let (A,B) be the pair (3.53). Let us prove that it is structurally stable.

For each matrix pair (C,D), we denote by $\mathcal{B}(C,D)$ its *bundle*, which is defined as follows (see [24]). Let

$$\mathcal{J}_{r_1}(\lambda_1) \oplus \cdots \oplus \mathcal{J}_{r_k}(\lambda_k) \oplus \mathcal{L}, \quad \lambda_1, \dots, \lambda_k \in \mathbb{C} \cup \infty,$$

be the Kronecker canonical form of (C,D), in which $\mathcal{J}_r(\lambda) := (I_r, J_r(\lambda))$ if $\lambda \in \mathbb{C}$, $\mathcal{J}_r(\infty) := (J_r(0), I_r)$, and $\mathcal{L}$ is a direct sum of pairs of the form (L_r, R_r) and (L_r^T, R_r^T). Then $\mathcal{B}(C,D)$ is the set of all matrix pairs with Kronecker canonical forms

$$\mathcal{J}_{r_1}(f(\lambda_1)) \oplus \cdots \oplus \mathcal{J}_{r_k}(f(\lambda_k)) \oplus \mathcal{L}, \quad f : \mathbb{C} \cup \infty \to \mathbb{C} \cup \infty \text{ bijective.}$$

In particular, $\mathcal{B} := \mathcal{B}(A,B)$ consists of all matrix pairs that are equivalent to (A,B).

Aiming for a contradiction, we suppose that each neighborhood of (A,B) has a pair that is not contained in $\mathcal{B}$. Since the set of matrix pairs of the same size is divided into a finite set of bundles, there is a bundle $\mathcal{C} \neq \mathcal{B}$ such that each neighborhood of (A,B) contains a pair from $\mathcal{C}$. This is impossible for the bundle $\mathcal{B}$ by [24, Theorem 3.3], in which the inclusion relationships between the closures of two bundles are described.

Therefore, there is a neighborhood of (A,B) that is contained in $\mathcal{B}$, which consists of the pairs that are equivalent to (A,B). $\square$

3.4.2. Stable matrix pairs under contragredient equivalence Each matrix pair in this section is a pair of complex matrices of sizes $m \times n$ and $n \times m$, in which $m, n = 1, 2, \dots$. We consider them up to transformations of *contragredient equivalence*

$$(A,B) \mapsto (S^{-1}AR, R^{-1}BS), \qquad S \text{ and } R \text{ are nonsingular.} \tag{3.54}$$

The matrices of a pair of counter linear mappings $U \rightleftarrows V$ are reduced by these transformations.

Dobrovol′skaya and Ponomarev [22] give a canonical form of a matrix pair under contragredient equivalence: each pair (A,B) is contragrediently equivalent to a direct sum, uniquely determined up to permutation of summands, of pairs of the types

$$(I_r, J_r(\lambda)),\ (J_r(0), I_r),\ (L_r, R_r^T),\ (L_r^T, R_r), \quad r = 1, 2, \dots, \tag{3.55}$$

in which L_r and R_r are defined in (3.52). Another proof of this canonical form and many applications are given by Horn and Merino [45].

Theorem 3.11. *Each pair (A,B) of complex matrices of sizes $m \times n$ and $n \times m$ with $m \geqslant 1$ and $n \geqslant 1$ is structurally unstable with respect to contragredient equivalence.*

Proof. It suffices to show that the pairs (3.55) are structurally unstable. If (A,B) and (A',B') are contragrediently equivalent, then AB and $A'B'$

are similar. Denote by E_{pq} the $p \times q$ matrix in which the (1,1) entry is an arbitrarily small complex number $\varepsilon \neq 0$ and the other entries are zero.

The pairs $(I_r, J_r(\lambda))$ and $(I_r, J_r(\lambda) + E_{rr})$ are not contragrediently equivalent since the matrices $I_r \cdot J_r(\lambda) = J_r(\lambda)$ and $I_r(J_r(\lambda) + E_{rr}) = J_r(\lambda) + E_{rr}$ are not similar.

The pairs (L_r, R_r^T) and $(L_r, R_r^T + E_{r,r-1})$ are not contragrediently equivalent since the matrices $L_r R_r^T = J_{r-1}(0)$ and $L_r(R_r^T + E_{r,r-1}) = J_{r-1}(0) + E_{r-1,r-1}$ are not similar.

Therefore, all pairs (3.55) are structurally unstable. $\square$

CONCLUSIONS AND FUTURE RESEARCH

The main goal of this thesis is to give a direct and constructive proof of Pokrzywa's theorem [60] that describes all possible Kronecker's canonical forms of matrix pencils in an arbitrary small neighborhood of a given matrix pencil.

More exactly, for each matrix pencil $A - xB$, Pokrzywa describes all Kronecker canonical forms $C - xD$ such that there exist arbitrarily small E and E' for which the Kronecker canonical form of $(A + E) - x(B + E')$ is $C - xD$. Pokrzywa's proof is very abstract and unconstructive; he does not construct $(A + E) - x(B + E')$ explicitly.

I construct $(A + E) - x(B + E')$ explicitly for each $C - xD$. As a consequence, I give a direct, constructive, and rather elementary proof of Pokrzywa's theorem.

For this purpose I calculate the Kronecker canonical forms of miniversal deformations of the following matrix pairs:

$$\begin{gathered}
(L_m, L_m) \oplus (L_n, R_n), \quad (L_m, R_m) \oplus (I_n, J_n(0)), \\
(L_m, R_m) \oplus (R_n^T, L_n^T), \quad (I_m, J_m(0)) \oplus (L_n, R_n), \\
(I_m, J_m(0)) \oplus (I_n, J_n(0)), \quad (I_m, J_m(0)) \oplus (L_n^T, L_n^T), \\
(R_m^T, L_m^T) \oplus (L_n, R_n), \quad (L_m^T, R_m^T) \oplus (I_n, J_n(0)), \\
(L_m^T, R_m^T) \oplus (L_n^T, R_n^T), \quad (I, J(\lambda)), \quad (J(0), I).
\end{gathered}$$

in which $J(\lambda)$ is a Jordan matrix with a single eigenvalue, and $J(0)$ is a nilpotent Jordan matrix.

In my thesis I also provide applications of perturbation theory which have been presented in 4 articles:

- *T. Klymchuk, Regularizing algorithm for mixed matrix pencils, Applied Mathematics and Nonlinear Sciences 2 (2017) 123–130;*
- *M. Isabel García-Planas, T. Klymchuk, Perturbation analysis of the matrix differential equation $\dot{x} = ABx$, Applied Mathematics and Nonlinear Sciences 3(1) (2018) 97–103;*
- *M. Isabel García-Planas, T. Klymchuk, Versal deformations of product of matrices under contragredient equivalence submitted to Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas, 2019; DOI: 10.1007/s13398-019-00678-5*
- *M. Isabel García-Planas, T. Klymchuk, Structural stability of matrix pencils and of matrix pairs under contragredient equivalence, Ukr. Mat. Zh. 71 (15) (2019) 706–709.*

Stratification is one of the fundamental concepts in algebraic geometry and the theory of singularities, thus in my future research I am going to continue the investigation of a stratification theory of matrix pencils under equivalence or contragredient equivalence using their miniversal deformations. More precisely I am going to apply obtained results to investigate a structural stability of matrix pencils under equivalence or contragredient equivalence and I am going to try to apply obtained result for Whitney stratification.

The study of stratifications originated with the work of Whitney and Thom [68,72]. This theory was developed by C.G. Gibson, K. Wirthmuller, A.A. Du Plessis, E.J.N. Looijenga in [42, 53], Alexandru Dimca in [18], J. Mather in [58, 59] and many others (see, for example, [66, 69, 73]). The translation of results obtained in this thesis to Whitney stratification is a very interesting and open problem which may have various applications in Algebraic Geometry and Catastrophe Theory.

Bibliography

[1] A. Alazemi, M. Andjelić, C. M. da Fonseca, V. V. Sergeichuk, *Lipschitz property for systems of linear mappings and bilinear forms*, Linear Algebra Appl. 573 (2019) 26–36.

[2] A. A. Andronov, L. Pontryagin, *Systèmes grossières*, Dokl. Akad. Nauk. SSR 14 (1937).

[3] V. I. Arnold, *On matrices depending on parameters*, Russian Math. Surveys 26 (2) (1971) 29–43.

[4] V. I. Arnold, *Lectures on bifurcations in versal families*, Russian Math Surveys 27 (5) (1972) 54–123.

[5] V. I. Arnold, *Geometrical Methods in the Theory of Ordinary Differential Equations*, Springer-Verlag, 1988.

[6] J. Berg, H. Kwantny, *A canonical parametrization of the Kronecker form of a matrix pencil*, Automatica 31 (1995) 669–680.

[7] F. M. Bleher, T. Chinburg, *Applications of versal deformations to Galois theory*, Commentarii Mathematici Helvetici 78(1) (2003) 45–64.

[8] K. Bongartz, *On degenerations and extensions of finite dimensional modules*, Adv. Math. 121 (2) (1996) 245–287.

[9] D. Boley, *Estimating the Sensitivity of the Algebraic Structure of Pencils with Simple Eigenvalue Estimates*, SIAM J. Matrix Anal. And Appl. 11(4) (1990) 632–643.

[10] D. Boley, *The algebraic structure of pencils and block Toeplitz matrices*, Linear Algebra Appl. 279 (1998) 255–279.

[11] S. Bora, M. Karow, C. Mehl, P. Sharma, *Structured eigenvalue backward errors of matrix pencils and polynomials with palindromic structures*, Siam J. Matrix Anal. Appl. 36(2) (2015) 393–416.

[12] V. Bolotnikov, *On the Sylvester equation over quaternions*, Noncommutative Analysis, Operator Theory and Applications (2016) 43–75.

[13] H. den Boer, G. Ph. A. Thijsse, *Semi-stability of sums of partial multiplicities under additive perturbation*, Integral Equations Operator Theory 3 (1980) 23–42.

[14] B. Conrad, A. J. De Jong, *Approximation of versal deformations*, Journal of Algebra 255(2) (2002) 489–515.

[15] D. D. de Oliveira, R. A. Horn, T. Klimchuk, V. V. Sergeichuk, *Remarks on the classification of a pair of commuting semilinear operators*, Linear Algebra Appl. 436 (2012) 3362–3372.

[16] J. Demmel and A. Edelman, *The dimension of matrices (matrix pencils) with given Jordan (Kronecker) canonical forms*, Linear Algebra Appl. 230 (1995) 61–87.

[17] D. Ž. Djoković, *Classification of pairs consisting of a linear and a semilinear map*, Linear Algebra Appl. 20 (1978) 147–165.

[18] Alexandru Dimca, *Singularities and Topology of Hypersurfaces*, Springer-Verlag New York, 1992.

[19] A. Dmytryshyn, F. M. Dopico, *Generic skew-symmetric matrix polynomials with fixed rank and fixed odd grade*, Linear Algebra Appl. 536 (2018) 1–18.

[20] A. Dmytryshyn, V. Futorny, V. V. Sergeichuk, *Miniversal deformations of matrices under *congruence and reducing transformations*, Linear Algebra and its Applications 446 (2014) 388–420.

[21] A. Dmytryshyn, B. Kågström, *Coupled Sylvester-type matrix equations and block diagonalization*, SIAM J. Matrix Anal. Appl. 36 (2015) 580–593.

[22] N. M. Dobrovol′skaya and V. A. Ponomarev, *A pair of counter-operators (in Russian)*, Uspehi Mat. Nauk 20 (6) (1965) 80–86.

[23] 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 (3) (1997) 653–692.

[24] 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.

[25] 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.

[26] J. Ferrer, M. I. García-Planas, F. Puerta, *Brunowsky local form of a holomorphic family of pairs of matrices*, Linear Algebra Appl. 253 (1997) 175–198.

[27] J. Ferrer, M. I. García-Planas, F. Puerta, *Regularity of the Brunovsky – Kronecker Stratification* SIAM Journal on Matrix Analysis and Applications 3(21) (2000) 724–742.

[28] P. Gabriel, A. V. Roiter, *Representations of Finite-Dimensional Algebras*, Encyclopaedia of Mathematical Sciences (7) 1992.

[29] D. M. Galin, *Real matrices depending on parameters*, Uspekhi Mat. Nauk 27:1 (163) (1972) 241–242.

[30] F. R. Gantmacher, *The Theory of Matrices*, Chelsea, New York, 1971.

[31] M. I. García-Planas, *Sensivity and stability of singular systems under proportional and derivative feedback*, WSEAS transactions on mathematics 8(11) (2009) 635–644.

[32] M. I. García-Planas, T. Klymchuk, *Perturbation analysis of a matrix differential equation $\dot{x} = ABx$*, Applied Mathematics and Nonlinear Sciences 3(1) (2018) 97-103.

[33] M. I. García-Planas, T. Klymchuk, *Versal deformations of matrix products*, Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas, 2019. DOI: 10.1007/s13398-019-00678-5

[34] M. I. García-Planas, T. Klymchuk, *Structural stability of matrix pencils and of matrix pairs under contragredient equivalence*, Ukr. Mat. Zh. 71 (15) (2019) 706–709.

[35] M. I. García-Planas, A. A. Mailybaev, *Reduction to versal deformations of matrix pencils and matrix pairs with application to control theory*, SIAM J. Matrix Anal. Appl. 24 (4) (2003) 943–962.

[36] M. I. García-Planas, M. D. Magret, *Deformation and stability of triples of matrices*, Linear Algebra Appl. 254 (1997) 159-192.

[37] M. I. García-Planas, M. D. Magret, *A Generalized Sylvester Equation: A Criterion for Structural Stability of Triples of Matrices*, Linear and Multilinear Algebra 44(2) (1998) 93–110.

[38] M. I. García-Planas, M. D. Magret, *Stratification of linear systems* Bifurcation diagrams for families of linear systems Linear algebra and its applications 297 (1999) 23–56.

[39] M. I. García-Planas, V. V. Sergeichuk, *Simplest miniversal deformations of matrices, matrix pencils, and contragredient matrix pencils*, Linear Algebra Appl. 302/303 (1999) 45–61.

[40] M. I. García-Planas, V. V. Sergeichuk, *Generic families of matrix pencils and their bifurcation diagrams*, Linear Algebra Appl. 332/334 (2001) 165–179.

[41] M. I. García-Planas, M. D. Magret, V. V. Sergeichuk, N. A. Zharko, *Rigid systems of second-order linear differential equations*, Linear Algebra Appl. 414 (2006) 517–532.

[42] C. G. Gibson, K. Wirthmüller, A. A. Du Plessis, E. J. N. Looijenga, *Topological Stability of Smooth Mappings*, Lecture Notes in Mathematics, Springer-Verlag (552) 1976.

[43] Y. P. Hong, R. A. Horn *A canonical form for matrices under consimilarity* Linear Algebra Appl. 102 (1988) 143–168.

[44] R. A. Horn, C. R. Johnson, *Matrix Analysis, 2nd ed.*, Cambridge University Press, Cambridge, 2013.

[45] R. A. Horn, D. I. Merino, *Contragredient equivalence: a canonical form and some applications*, Linear Algebra Appl. 214 (1995) 43–92.

[46] R. A. Horn, V. V. Sergeichuk, *A regularization algorithm for matrices of bilinear and sesquilinear forms*, Linear Algebra Appl. 412 (2006) 380–395.

[47] P. Johansson, *StratiGraph User's 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

[48] L. Klimenko, V. V. Sergeichuk, *Block triangular miniversal deformations of matrices and matrix pencils*, Matrix Methods: Theory, Algorithms and Applications, World Sci. Publ., Hackensack, NJ (2010) 69–84.

[49] H. Koçak, *Normal forms and versal deformations of linear Hamiltonian systems*, Journal of Differential Equations 51(3) (1984) 359–407.

[50] D. Kressner, Xin Liu, *Structured canonical forms for products of (skew-) symmetric matrices and the matrix equation $XAX = B$*, Electronic Journal of Linear Algebra 26 (2013) 215–230.

[51] T. Klymchuk, *Regularizing algorithm for mixed matrix pencils*, Applied Mathematics and Nonlinear Sciences 2 (2017) 123–130.

[52] P. Lancaster, M. Tismenetsky, *The Theory of Matrices with Applications, 2nd ed.*, Academic Press, 1985.

[53] E. Looijenga, *Structural stability of families of C^∞-functions and the canonical stratification of $C^\infty(N)$*, I.H.E.S. Math. Jan, 1974.

[54] S. López de Medrano, *Topological aspects of matrix problems, in: Representations of algebras (Puebla, 1980)*, Lecture Notes in Math. 903, Springer (1981) 96–210.

[55] 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 (electronic).

[56] A. A. Mailybaev, *Transformation to versal deformations of matrices*, Linear Algebra Appl. 337 (2001) 87–108.

[57] 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.

[58] J. N. Mather, *Stratification and Mappings, Dynamical Systems*, Academic Press, 1973.

[59] J. N. Mather, *How to stratify mappings and jet space*, (to appear).

[60] A. Pokrzywa, *On perturbations and the equivalence orbit of a matrix pencil*, Linear Algebra Appl. 82 (1986) 99–121.

[61] L. Rodman, *Remarks on Lipschitz properties of matrix groups actions*, Linear Algebra Appl. 434 (2011) 1513–1524.

[62] J. W. Robbin, *Topological conjugacy and structural stability for discrete dynamical systems*, Bull. Amer. Math. Soc. 78 (1972) 923–952.

[63] P. Rubio, J. Gelonch, *The contragredient equivalence: Application to solve some matrix systems*, Linear Algebra and its Applications 290 (1999) 145–166.

[64] V. V. Sergeichuk, *Canonical matrices for linear matrix problems*, Linear Algebra Appl. 317(1-3) (2000) 53–102.

[65] V. V. Sergeichuk, *Computation of canonical matrices for chains and cycles of linear mappings*, Linear Algebra Appl. 376 (2002) 235–263.

[66] J. P. Speder, *Equisingularite et conditions de Whitney*, Nice I.M.S.P. 1972–1973.

[67] F. de Terán, F. M. Dopico *A note of generic Kronecker orbits of matrix pencils with fixed rank*, SIAM J. Matrix Anal. Appl. 30(2) (2008) 491–496.

[68] R. Thom, *Ensemble et morphisms stratifies*, Bulletin AMS 75 (1969) 240–284.

[69] G. N. Tyurina, *Resolution of singularities of plane deformations of double rational points*, Functional Anal. Appl. 4 (1970) 68–73.

[70] P. Van Dooren, *The computation of Kronecker's canonical form of a singular pencil*, Linear Algebra Appl. 27 (1979) 103–140.

[71] A. Varga, *Computation of Kronecker-like forms of periodic matrix pairs*, Symp. on Mathematical Theory of Networks and Systems, Leuven, Belgium, July 5–9, 2004.

[72] H. Whitney, *Elementary Structure of Real Algebraic Varieties*, Ann. Math. 66 (1957) 545–556.

[73] H. Whitney, *Tangents to an Analytic Variety*, Ann. of Math. 81 (1965) 496–540.

[74] J. C. Willems, *Topological classification and structural stability of linear systems*, J. Differential Equations 35 (3) (1980) 306–318.

[75] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, 1965.