ARTICLE IN PRESS LAA:15660

Linear Algebra and its Applications ••• (••••) •••–•••

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Perturbation theory of matrix pencils through miniversal deformations

Vyacheslav Futorny [a], Tetiana Klymchuk [b], Olena Klymenko [c], Vladimir V. Sergeichuk [d,*], Nadiya Shvai [e]

[a] *Department of Mathematics, University of São Paulo, Brazil*
[b] *International University of Catalonia, Barcelona, Spain*
[c] *National Technical University of Ukraine "Kyiv Polytechnic Institute", Kyiv, Ukraine*
[d] *Institute of Mathematics, Kyiv, Ukraine*
[e] *National University of Kyiv-Mohyla Academy, Kyiv, Ukraine*

A R T I C L E I N F O

Article history:
Received 30 October 2019
Accepted 6 December 2020
Available online xxxx
Submitted by R. Brualdi

MSC:
15A21
15A22

Keywords:
Matrix pencils
Kronecker canonical form
Perturbations
Orbit closures

A B S T R A C T

We give new proofs of several known results about perturbations of matrix pencils. In particular, we give a direct and constructive proof of Andrzej Pokrzywa's theorem (1983), in which the closure of the orbit of each Kronecker canonical matrix pencil is described in terms of inequalities for invariants of matrix pencils. A more abstract description is given by Klaus Bongartz (1996) by methods of representation theory. We formulate and prove Pokrzywa's theorem in terms of successive replacements of direct summands in a Kronecker canonical pencil.

First we show that it is sufficient to prove Pokrzywa's theorem in two cases: for matrices under similarity and for each matrix pencil $P - \lambda Q$ that is a direct sum of two indecomposable pencils. Then we calculate the Kronecker canonical form of pencils that are close to $P - \lambda Q$. In fact, the Kronecker canonical form is calculated for only those pencils that belong to a miniversal deformation of $P - \lambda Q$. This is sufficient since

* Corresponding author.
E-mail addresses: futorny@ime.usp.br (V. Futorny), klimchuk.tanya@gmail.com (T. Klymchuk), e.n.klimenko@gmail.com (O. Klymenko), sergeich@imath.kiev.ua (V.V. Sergeichuk), nadiia.shvai@gmail.com (N. Shvai).

https://doi.org/10.1016/j.laa.2020.12.009

Please cite this article in press as:, Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

all pencils in a neighborhood of $P - \lambda Q$ are reduced to them by a smooth strict equivalence transformation.

1. Introduction

For each complex Jordan matrix A, Den Boer and Thijsse [13] and, independently, Markus and Parilis [35] describe all Jordan matrices J such that each neighborhood of A contains a matrix that is similar to J. Pokrzywa [37] extends their results to Kronecker canonical pencils $A - \lambda B$ ($A, B \in \mathbb{C}^{m \times n}$): he describes the set of all Kronecker canonical pencils $K - \lambda L$ such that each neighborhood of $A - \lambda B$ contains a pencil whose Kronecker canonical form is $K - \lambda L$. Pokrzywa formulates and proves his theorem in terms of inequalities for invariants of matrix pencils. A more abstract solution of this problem is given by Bongartz [9, Section 5, Table I] by methods of representation theory (see also [5,7,8,10]).

The main purpose of this paper is to give a direct and constructive proof of Pokrzywa's theorem using García-Planas and Sergeichuk's miniversal deformations of matrix pencils [27].

Instead of pencils $A - \lambda B$, we consider matrix pairs (A, B). We study them up to *equivalence transformations*

$$(A, B) \mapsto (SAR, SBR), \qquad S \text{ and } R \text{ are nonsingular matrices.}$$

For each pair $\mathcal{A} = (A, B)$, its *orbit* $\langle \mathcal{A} \rangle$ is the set of all pairs that are equivalent to $\mathcal{A}$.

Let $\mathcal{P}_{m,n}$ be the set of orbits of pairs of $m \times n$ complex matrices. Pokrzywa's theorem describes the following partial ordering on $\mathcal{P}_{m,n}$: $\langle \mathcal{A} \rangle \leqslant \langle \mathcal{B} \rangle$ if and only if $\langle \mathcal{A} \rangle$ is contained in the closure of $\langle \mathcal{B} \rangle$. Thus,

$$\begin{gathered}\langle \mathcal{A} \rangle \leqslant \langle \mathcal{B} \rangle \text{ if and only if } \mathcal{A} \text{ can be transformed by an arbitrarily small} \\ \text{perturbation to a pair that is equivalent to } \mathcal{B}.\end{gathered} \tag{1}$$

An orbit $\langle \mathcal{B} \rangle$ *immediately succeeds* $\langle \mathcal{A} \rangle$ (many authors write that $\langle \mathcal{B} \rangle$ *covers* $\langle \mathcal{A} \rangle$; see [22]) if $\langle \mathcal{A} \rangle < \langle \mathcal{B} \rangle$ and there exists no $\langle \mathcal{C} \rangle$ such that $\langle \mathcal{A} \rangle < \langle \mathcal{C} \rangle < \langle \mathcal{B} \rangle$.

The partially ordered set $\mathcal{P}_{m,n}$ is visually represented by its *Hasse diagram* (also called the closure graph), which is the directed graph whose vertices are the orbits from $\mathcal{P}_{m,n}$ and there is an arrow $\langle \mathcal{A} \rangle \longrightarrow \langle \mathcal{B} \rangle$ if and only if $\langle \mathcal{B} \rangle$ immediately succeeds $\langle \mathcal{A} \rangle$.

Example 1.1. Each pair of 1×2 matrices is equivalent to exactly one of the pairs

$$([0\ 0], [0\ 0]), \quad ([1\ 0], [\lambda\ 0]) \text{ with } \lambda \in \mathbb{C}, \quad ([0\ 0], [1\ 0]), \quad ([1\ 0], [0\ 1])$$

(they are the Kronecker canonical pairs of 1×2 matrices for equivalence; see (4)). The Hasse diagram of $\mathcal{P}_{1,2}$ is

$$\begin{tikzcd}[column sep=small]
& \langle([1\ 0],[0\ 1])\rangle & \\
\langle([1\ 0],[\lambda\ 0])\rangle \arrow[ur] & & \langle([0\ 0],[1\ 0])\rangle \arrow[ul] \\
& \langle([0\ 0],[0\ 0])\rangle \arrow[ul] \arrow[ur] &
\end{tikzcd} \tag{2}$$

By (1), for each arrow $\langle\mathcal{A}\rangle \longrightarrow \langle\mathcal{B}\rangle$ there exists an arbitrarily small perturbation $\Delta\mathcal{A}$ such that $\mathcal{A} + \Delta\mathcal{A}$ is equivalent to $\mathcal{B}$; we position $\Delta\mathcal{A}$ on the corresponding arrow of (2):

$$\begin{tikzcd}[column sep=small]
& \langle([1\ 0],[0\ 1])\rangle & \\
\langle([1\ 0],[\lambda\ 0])\rangle \arrow[ur, "{([0\ \ 0],[0\ \ \varepsilon])}" description] & & \langle([0\ 0],[1\ 0])\rangle \arrow[ul, "{([0\ \ \varepsilon],[0\ \ 0])}" description] \\
& \langle([0\ 0],[0\ 0])\rangle \arrow[ul, "{([\varepsilon\ \ 0],[\varepsilon\lambda\ \ 0])}" description] \arrow[ur, "{([0\ \ 0],[\varepsilon\ \ 0])}" description] &
\end{tikzcd}$$

in which ε is an arbitrarily small complex number.

The Hasse diagram of $\mathcal{P}_{2,3}$ is given in [24]. The Hasse diagram of $\mathcal{P}_{m,n}$ with arbitrary m and n is constructed by the software StratiGraph [23,29,40], which is based on Pokrzywa's theorem. The Hasse diagrams for congruence classes of 2×2 and 3×3 complex matrices and for *congruence classes of 2×2 complex matrices are constructed in [18,25]. The Hasse diagrams for matrix polynomials are constructed in [20].

The main theorem of the paper is *Theorem I* from Section 2, which is another form of Pokrzywa's theorem. Theorem I gives six types of replacements of direct summands such that a Kronecker pair $\mathcal{A}$ is transformed to a Kronecker pair $\mathcal{B}$ by a sequence of replacements of these types if and only if $\langle\mathcal{A}\rangle < \langle\mathcal{B}\rangle$. Two principal tools in our proof of Theorem I are the following:

(a) *Theorem 4.1*, which states that each immediate successor of the orbit of a Kronecker pair $\mathcal{A}$ is the orbit of a pair that is obtained from $\mathcal{A}$ by an arbitrarily small perturbation of only one of its subpairs of the form
  (i) (P, Q), which is an indecomposable Kronecker pair (see (4)) or the direct sum of two indecomposable Kronecker pairs, and
  (ii) (I, J), in which J is a Jordan matrix with single eigenvalue.

Please cite this article in press as: V. Futorny et al., Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

Thus, it is sufficient to prove Theorem I for (P, Q) and (I, J); we do this in Sections 5 and 6.

(b) *García-Planas and Sergeichuk's miniversal deformations* of matrix pairs under equivalence; they are given in [27] and are presented in Section 3. In Section 5, we calculate the Kronecker canonical form of pairs that are close to a pair (P, Q) from (i). In fact, we calculate the Kronecker canonical form of only those pairs that belong to the miniversal deformation of (P, Q) given in [27], which is sufficient since all pairs close to (P, Q) are reduced to such pairs by smooth equivalence transformations. This simplifies the calculation cardinally since miniversal perturbations do not change many entries. For example, all matrices that are close to the Jordan block $J_3(\lambda)$ and those of them that form a miniversal deformation of $J_3(\lambda)$ are of the form

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

respectively, in which the stars are complex numbers.

A miniversal deformation of each Kronecker canonical pair (A, B) was first constructed by Edelman, Elmroth, and Kågström [21]. We use García-Planas and Sergeichuk's miniversal deformation given in [27] since it is simpler: it consists of pairs of the form $(A+X, B+Y)$, in which all nonzero entries of X and Y are non-repeating independent parameters. For example the miniversal deformations of $(I_3, J_3(\lambda))$ in [21] and [27] are families of matrix pairs

$$\left(I_3,\begin{bmatrix}\lambda&1&0\\0&\lambda&1\\0&0&\lambda\end{bmatrix}+\begin{bmatrix}a&0&0\\b&a&0\\c&b&a\end{bmatrix}\right)\quad\text{and}\quad\left(I_3,\begin{bmatrix}\lambda&1&0\\0&\lambda&1\\0&0&\lambda\end{bmatrix}+\begin{bmatrix}0&0&0\\0&0&0\\c&b&a\end{bmatrix}\right),\tag{3}$$

in which $a, b, c \in \mathbb{C}$.

The article is organized as follows. In Section 2 we formulate Theorem I about sequences of replacements that transform a Kronecker pair $\mathcal{A}$ to a Kronecker pair $\mathcal{B}$ such that $\langle\mathcal{A}\rangle < \langle\mathcal{B}\rangle$, and Theorem II that describes when $\mathcal{B}$ is an immediate successor of $\langle\mathcal{A}\rangle$. In Section 3 we recall miniversal deformations of matrix pairs under equivalence given in [27]. In Section 4 we reduce the proof of Theorem I to the case of pairs that are direct sums of two indecomposable pairs and to the case of pairs of the form (I, J), in which J is a Jordan matrix with a single eigenvalue. Theorem I is proved for these pairs in Sections 5 and 6. Theorem II is proved in Section 7.

2. Main theorems

All matrices that we consider are complex matrices and each matrix pair consists of matrices of the same size. For each positive integer n, we define the matrices[1]

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

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

We also define the matrices

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

whose sizes will be clear from the context.

The matrix pairs

$$\mathcal{L}_n := (L_n, R_n), \quad \mathcal{L}_n^T := (L_n^T, R_n^T), \quad \mathcal{D}_n(\lambda) := \begin{cases} (I_n, J_n(\lambda)) \text{ if } \lambda \in \mathbb{C} \\ (J_n(0), I_n) \text{ if } \lambda = \infty \end{cases} \tag{4}$$

are called *indecomposable Kronecker pairs*. Leopold Kronecker proved that each matrix pair $\mathcal{A}$ is equivalent to a direct sum of such pairs. This direct sum is called the *Kronecker canonical form* of $\mathcal{A}$; it is uniquely determined by $\mathcal{A}$, up to permutation of direct summands.

2.1. First main theorem

The closures of orbits of Kronecker pairs are described in the following theorem.

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 a sequence of permutations of direct summands and listed below replacements* (i)–(vi) *of direct summands, in which $m, n \in \{1, 2, \ldots\}$ and $\lambda \in \mathbb{C} \cup \infty$. The notation $\mathcal{P} \rightsquigarrow \mathcal{Q}$ means that $\mathcal{P}$ is replaced by $\mathcal{Q}$. For each replacement $\mathcal{P} \rightsquigarrow \mathcal{Q}$, we also give a pair that is obtained from $\mathcal{P}$ by an arbitrarily small perturbation (which is defined by an arbitrary nonzero complex number ε) and whose Kronecker canonical form is $\mathcal{Q}$.*

[1] For each nonnegative integers p and q, we denote by 0_{pq} the zero matrix of size $p \times q$. In particular, $L_1 = R_1 = 0_{01}$. If M is an $m \times n$ matrix, then $M \oplus 0_{0q} = [\,M \quad 0_{mq}\,]$ and $M \oplus 0_{p0} = \begin{bmatrix} M \\ 0_{pn} \end{bmatrix}$.

Please cite this article in press as: V. Futorny et al., Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

(i) $\mathcal{L}_m^T \oplus \mathcal{L}_n^T \rightsquigarrow \mathcal{L}_{m+1}^T \oplus \mathcal{L}_{n-1}^T$ *in which* $m+2 \leqslant n$, *via the pair*

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

which is obtained by perturbation of $\mathcal{L}_m^T \oplus \mathcal{L}_n^T$.

(ii) $\mathcal{L}_m \oplus \mathcal{L}_n \rightsquigarrow \mathcal{L}_{m+1} \oplus \mathcal{L}_{n-1}$ *in which* $m+2 \leqslant n$, *via*

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

(iii) $\mathcal{L}_m^T \oplus \mathcal{D}_n(\lambda) \rightsquigarrow \mathcal{L}_{m+1}^T \oplus \mathcal{D}_{n-1}(\lambda)$ *(if* $n = 1$, *then the summand* $\mathcal{D}_0(\lambda)$ *is omitted), via*

$$\left(\begin{bmatrix} L_m^T & 0 \\ 0 & I_n \end{bmatrix}, \begin{bmatrix} R_m^T & \varepsilon 0^{\nearrow} \\ 0 & J_n(\lambda) \end{bmatrix}\right) \text{ if } \lambda \in \mathbb{C}, \quad \left(\begin{bmatrix} L_m^T & \varepsilon 0^{\searrow} \\ 0 & J_n(0) \end{bmatrix}, \begin{bmatrix} R_m^T & 0 \\ 0 & I_n \end{bmatrix}\right) \text{ if } \lambda = \infty.$$

(iv) $\mathcal{L}_m \oplus \mathcal{D}_n(\lambda) \rightsquigarrow \mathcal{L}_{m+1} \oplus \mathcal{D}_{n-1}(\lambda)$, *via*

$$\left(\begin{bmatrix} L_m & 0 \\ 0 & I_n \end{bmatrix}, \begin{bmatrix} R_m & 0 \\ \varepsilon 0^{\nwarrow} & J_n(\lambda) \end{bmatrix}\right) \text{ if } \lambda \in \mathbb{C}, \quad \left(\begin{bmatrix} L_m & 0 \\ \varepsilon 0^{\nearrow} & J_n(0) \end{bmatrix}, \begin{bmatrix} R_m & 0 \\ 0 & I_n \end{bmatrix}\right) \text{ if } \lambda = \infty.$$

(v) $\mathcal{D}_m(\lambda) \oplus \mathcal{D}_n(\lambda) \rightsquigarrow \mathcal{D}_{m-1}(\lambda) \oplus \mathcal{D}_{n+1}(\lambda)$ *in which* $m \leqslant n$, *via*

$$\left(I_{m+n}, \begin{bmatrix} J_m(\lambda) & \varepsilon 0^{\nwarrow} \\ 0 & J_n(\lambda) \end{bmatrix}\right) \text{ if } \lambda \in \mathbb{C}, \quad \left(\begin{bmatrix} J_m(0) & \varepsilon 0^{\nwarrow} \\ 0 & J_n(0) \end{bmatrix}, I_{m+n}\right) \text{ if } \lambda = \infty.$$

(vi) $\mathcal{L}_m^T \oplus \mathcal{L}_n \rightsquigarrow \mathcal{D}_{r_1}(\mu_1) \oplus \cdots \oplus \mathcal{D}_{r_k}(\mu_k)$, *in which* $\mu_1, \ldots, \mu_k \in \mathbb{C} \cup \infty$ *are distinct and* $r_1 + \cdots + r_k = m+n-1$, *via the pair*

$$\left(\left[\begin{array}{cccc|cccc} 1 & & & & & & & \alpha_1 \\ 0 & \ddots & & & & & & \alpha_2 \\ & \ddots & 1 & & & 0 & & \vdots \\ & & 0 & & & & & \alpha_m \\ \hline & & & & 1 & 0 & & \\ & 0 & & & & \ddots & \ddots & \\ & & & & & & 1 & 0 \end{array}\right], \left[\begin{array}{cccc|cccc} 0 & & & & \beta_1 & \beta_2 & \ldots & \beta_n \\ 1 & \ddots & & & & & & \\ & \ddots & 0 & & & 0 & & \\ & & 1 & & & & & \\ \hline & & & & 0 & 1 & & \\ & 0 & & & & \ddots & \ddots & \\ & & & & & & 0 & 1 \end{array}\right]\right) \tag{5}$$

that is defined by

$$\prod_{\mu_i \neq \infty} (x - \mu_i)^{r_i} = c_0 + c_1 x + \cdots + c_{r-1} x^{r-1} + x^r, \tag{6}$$

$$\varepsilon(c_0, \ldots, c_{r-1}, 1, 0, \ldots, 0) = (-\beta_1, \ldots, -\beta_n, \alpha_1, \ldots, \alpha_m), \tag{7}$$

in which ε is any nonzero complex number.

Pokrzywa [37, Theorem 3] describes the closures of orbits of Kronecker canonical pencils in the form of systems of inequalities for invariants of matrix pencils (see also [12, Theorem 2.1] and [22, Theorem 3.1]). Nevertheless, he formulates Lemma 5 in [37] in the form of replacements of direct summands of Kronecker pairs; such replacements are also given in [6, Section 5.1] and [16, Theorem 2.2]. In the proof of Lemma 5 in [37], Pokrzywa also gives arbitrarily small perturbations that produce these replacements.

We will show that the statements (i)–(vi) of Theorem I follow from Theorems 5.1–5.6 by Theorems 4.1 and 6.1.

2.2. Second main theorem

The following theorem was first given by Edelman, Elmroth, and Kågström [22, Theorem 3.2] in the form of coin moves; see also [28, Theorem 2.4] and [5]. We derive it from Theorem I in Section 7, which can be read independently of Sections 3–6.

Theorem II. *Let $\mathcal{A}$ be a Kronecker pair. An orbit $\mathcal{O}$ immediately succeeds $\langle\mathcal{A}\rangle$ if and only if $\mathcal{O}$ is the orbit of a pair that is obtained from $\mathcal{A}$ by exactly one of the following replacements, which are special cases of the replacements* (i)–(vi) *of Theorem I:*

- (i′) $\mathcal{L}_m^T \oplus \mathcal{L}_n^T \rightsquigarrow \mathcal{L}_{m+1}^T \oplus \mathcal{L}_{n-1}^T$ $(m+2 \leqslant n)$ *such that if $\mathcal{A}$ contains $\mathcal{L}_m^T \oplus \mathcal{L}_k^T \oplus \mathcal{L}_n^T$ with $m<k<n$, then $n-m=2$,*
- (ii′) $\mathcal{L}_m \oplus \mathcal{L}_n \rightsquigarrow \mathcal{L}_{m+1} \oplus \mathcal{L}_{n-1}$ $(m+2 \leqslant n)$ *such that if $\mathcal{A}$ contains $\mathcal{L}_m \oplus \mathcal{L}_k \oplus \mathcal{L}_n$ with $m<k<n$, then $n-m=2$,*
- (iii′) $\mathcal{L}_{\overline{m}}^T \oplus \mathcal{D}_{\overline{n}_\lambda}(\lambda) \rightsquigarrow \mathcal{L}_{\overline{m}+1}^T \oplus \mathcal{D}_{\overline{n}_\lambda-1}(\lambda)$ *(if $\mathcal{A}$ contains pairs of the form $\mathcal{L}_m^T$ and $\mathcal{D}_n(\lambda)$), in which $\overline{m} := \max\{m \,|\, \mathcal{L}_m^T$ in $\mathcal{A}\}$ and $\overline{n}_\lambda := \max\{n \,|\, \mathcal{D}_n(\lambda)$ in $\mathcal{A}\}$,*
- (iv′) $\mathcal{L}_{\overline{m}} \oplus \mathcal{D}_{\overline{n}_\lambda}(\lambda) \rightsquigarrow \mathcal{L}_{\overline{m}+1} \oplus \mathcal{D}_{\overline{n}_\lambda-1}(\lambda)$*, in which $\overline{m} := \max\{m \,|\, \mathcal{L}_m$ in $\mathcal{A}\}$ and $\overline{n}_\lambda := \max\{n \,|\, \mathcal{D}_n(\lambda)$ in $\mathcal{A}\}$,*
- (v′) $\mathcal{D}_m(\lambda) \oplus \mathcal{D}_n(\lambda) \rightsquigarrow \mathcal{D}_{m-1}(\lambda) \oplus \mathcal{D}_{n+1}(\lambda)$ $(m \leqslant n)$ *such that $\mathcal{A}$ does not contain $\mathcal{D}_m(\lambda) \oplus \mathcal{D}_k(\lambda) \oplus \mathcal{D}_n(\lambda)$ with $m \leqslant k \leqslant n$ and $m<n$,*
- (vi′) $\mathcal{L}_{\overline{m}}^T \oplus \mathcal{L}_{\overline{n}} \rightsquigarrow \mathcal{D}_{r_1}(\mu_1) \oplus \cdots \oplus \mathcal{D}_{r_k}(\mu_k)$ *($\mu_1,\dots,\mu_k \in \mathbb{C}\cup\infty$ are distinct and $r_1+\cdots+r_k = \overline{m}+\overline{n}-1$), in which $\overline{m} := \max\{m \,|\, \mathcal{L}_m^T$ in $\mathcal{A}\}$, $\overline{n} := \max\{n \,|\, \mathcal{L}_n$ in $\mathcal{A}\}$, and if $\mathcal{D}_k(\lambda)$ is contained in $\mathcal{A}$ then $\lambda=\mu_i$ for some i and $r_i \geqslant k$.*

Remark 2.1. Theorems I and II generalize the following known description (see [22, Section 2]) of the closures of orbits of Jordan matrices. Let J and J' be non-similar Jordan matrices.

- $\langle J\rangle < \langle \mathcal{J}'\rangle$ if and only if J' can be obtained from J by a sequence of permutations of Jordan blocks and replacements of pairs of direct summands

Please cite this article in press as: V. Futorny et al., Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

$$J_m(\lambda) \oplus J_n(\lambda)\ (\lambda \in \mathbb{C},\ m \leqslant n) \quad \text{by} \quad J_{m-1}(\lambda) \oplus J_{n+1}(\lambda). \tag{8}$$

- $\langle J' \rangle$ immediately succeeds $\langle J \rangle$ if and only if J' is obtained from J by a permutation of Jordan blocks and exactly one replacement (8) such that J does not contain $J_m(\lambda) \oplus J_k(\lambda) \oplus J_n(\lambda)$ with $m < n$ and $m \leqslant k \leqslant n$.

2.3. Third main theorem

Define the matrices whose sizes will be clear from the context:

$$\Delta_r(\varepsilon) := \begin{bmatrix} 0 \ldots 0\, \varepsilon\, 0 \ldots 0 \\ 0 \end{bmatrix}, \qquad \nabla_r(\varepsilon) := \begin{bmatrix} 0 \\ 0 \ldots 0\, \varepsilon\, 0 \ldots 0 \end{bmatrix}, \tag{9}$$

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 $\Delta_0 = \nabla_0 := 0$.

The *lower cone* of an orbit $\langle \mathcal{A} \rangle$ is the set $\langle \mathcal{A} \rangle^\vee$ of all orbits $\langle \mathcal{B} \rangle$ such that $\langle \mathcal{A} \rangle \leqslant \langle \mathcal{B} \rangle$. Theorems 5.1–5.6 are used in the proof of Theorem I; however, they are also imply the following theorem.

Theorem III. • *If $\mathcal{A}$ is an indecomposable Kronecker pair, then $\langle \mathcal{A} \rangle^\vee$ is the one-element set $\{\langle \mathcal{A} \rangle\}$.*

• *The direct sums of two indecomposable Kronecker pairs have the following lower cones (in which ε is an arbitrary nonzero complex number):*

(i) *The cone $\langle \mathcal{L}_m^T \oplus \mathcal{L}_n^T \rangle^\vee$ with $1 \leqslant m \leqslant n$ consists of the orbits of*

$$\mathcal{L}_{m+r}^T \oplus \mathcal{L}_{n-r}^T, \quad \text{in which } r \geqslant 0 \text{ and } m + r \leqslant n - r. \tag{10}$$

Each pair (10) is the Kronecker canonical form of

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

(ii) *$\langle \mathcal{L}_m \oplus \mathcal{L}_n \rangle^\vee$ with $1 \leqslant m \leqslant n$ consists of the orbits of*

$$\mathcal{L}_{m+r} \oplus \mathcal{L}_{n-r}, \quad \text{in which } r \geqslant 0 \text{ and } m + r \leqslant n - r. \tag{11}$$

Each pair (11) is the Kronecker canonical form of

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

(iii) *$\langle \mathcal{L}_m^T \oplus \mathcal{D}_n(\lambda) \rangle^\vee$ with $m \geqslant 1$, $n \geqslant 1$, and $\lambda \in \mathbb{C} \cup \infty$ consists of the orbits of*

$$\mathcal{L}_{m+r}^T \oplus \mathcal{D}_{n-r}(\lambda), \quad \text{in which } 0 \leqslant r \leqslant n. \tag{12}$$

Each pair (12) *is the Kronecker canonical form of*

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

(iv) $\langle \mathcal{L}_m \oplus \mathcal{D}_n(\lambda)\rangle^\vee$ *with* $m \geqslant 1$, $n \geqslant 1$, *and* $\lambda \in \mathbb{C} \cup \infty$ *consists of the orbits of*

$$\mathcal{L}_{m+r} \oplus \mathcal{D}_{n-r}(\lambda), \quad \text{in which } 0 \leqslant r \leqslant n. \tag{13}$$

Each pair (13) *is the Kronecker canonical form of*

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

(v) $\langle \mathcal{D}_m(\lambda) \oplus \mathcal{D}_n(\lambda)\rangle^\vee$ *with* $1 \leqslant m \leqslant n$ *and* $\lambda \in \mathbb{C} \cup \infty$ *consists of the orbits of*

$$\mathcal{D}_{m-r}(\lambda) \oplus \mathcal{D}_{n+r}(\lambda), \quad \text{in which } 0 \leqslant r \leqslant m. \tag{14}$$

Each pair (14) *is the Kronecker canonical form of*

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

(vi) $\langle \mathcal{L}_m^T \oplus \mathcal{L}_n\rangle^\vee$ *with* $m \geqslant 1$ *and* $n \geqslant 1$ *consists of* $\langle \mathcal{L}_m^T \oplus \mathcal{L}_n\rangle$ *and the orbits of*

$$\mathcal{D}_{r_1}(\mu_1) \oplus \cdots \oplus \mathcal{D}_{r_k}(\mu_k), \qquad \begin{matrix} r_1 + \cdots + r_k = m + n - 1, \\ \mu_1, \ldots, \mu_k \in \mathbb{C} \cup \infty \text{ are distinct.} \end{matrix} \tag{15}$$

Each pair (15) *is the Kronecker canonical form of the pair* (5) *determined by* (6) *and* (7).

2.4. Fourth main theorem

Let $\mathcal{A}$ be a Kronecker pair whose direct summands are arranged as follows:

$$\begin{aligned} \mathcal{A} = &\mathcal{L}_{m_1}^T \oplus \mathcal{L}_{m_2}^T \oplus \cdots \oplus \mathcal{L}_{m_{\underline{s}}}^T \\ &\oplus \bigoplus_{i=1}^{t} \Big(\mathcal{D}_{k_{i1}}(\lambda_i) \oplus \mathcal{D}_{k_{i2}}(\lambda_i) \oplus \cdots \oplus \mathcal{D}_{k_{is_i}}(\lambda_i) \Big) \\ &\oplus \mathcal{L}_{n_{\overline{s}}} \oplus \mathcal{L}_{n_{\overline{s}-1}} \oplus \cdots \oplus \mathcal{L}_{n_1}, \end{aligned} \tag{16}$$
$$m_1 \leqslant \cdots \leqslant m_{\underline{s}}, \quad k_{i1} \leqslant \cdots \leqslant k_{is_i} \ (i = 1, \ldots, t), \quad n_1 \leqslant \cdots \leqslant n_{\overline{s}}.$$

The numbers $\underline{s}, t, \overline{s}$ can be zero, which means that the corresponding direct summands in (16) are absent. By the following theorem, each immediate successor of $\langle \mathcal{A}\rangle$ is the orbit of a pair that is obtained by an arbitrarily small perturbation of only one pair of conformally located upper diagonal blocks of $\mathcal{A}$.

Please cite this article in press as: V. Futorny et al., Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

Theorem IV. *Let $\mathcal{A} = ([A_{ij}], [A'_{ij}])$ be a Kronecker pair of form (16) partitioned into blocks A_{ij} and A'_{ij} such that the pairs of diagonal blocks (A_{11}, A'_{11}), (A_{22}, A'_{22}), ... are the direct summands*

$$\begin{aligned} &\mathcal{L}^T_{m_1}, \dots, \mathcal{L}^T_{m_{\underline{s}}}, \ \ \mathcal{D}_{k_{11}}(\lambda_1), \dots, D_{k_{1s_1}}(\lambda_1), \\ &\qquad\qquad\qquad \dots, \ \ \mathcal{D}_{k_{t1}}(\lambda_t), \dots, D_{k_{ts_t}}(\lambda_t), \ \ \mathcal{L}_{n_{\overline{s}}}, \dots, \mathcal{L}_{n_1} \end{aligned} \tag{17}$$

of (16). 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.

Theorem I implies Theorem IV since all perturbations in (i)–(vi) applied to (16) are upper block-triangular. We move backwards in the next sections: we first give an independent proof of Theorem 4.1, which is a weak form of Theorem IV. Using it, we prove Theorem I in Sections 5 and 6.

3. Preliminaries

3.1. Miniversal deformations of matrices

Vladimir Arnold [2] defines a *deformation* of a square complex matrix A as a matrix $A(y_1, \dots, y_t)$ of the same size with entries that are power series of complex variables $y_1, \dots, y_t$ convergent in a neighborhood of $(0, \dots, 0)$ with $A(0, \dots, 0) = A$. A deformation is *linear* if its entries are linear polynomials:

$$A(y_1, \dots, y_t) = A + A_1 y_1 + \dots + A_t y_t, \qquad A, A_1, \dots, A_t \in \mathbb{C}^{n \times n}.$$

Arnold also considers a deformation $A(y_1, \dots, y_t)$ as a family of matrices with parameters $y_1, \dots, y_t$. If all matrices $A + X$ close to A can be reduced to matrices from this family by a similarity transformation $S(X)^{-1}(A + X)S(X)$ in which $S(X)$ is a deformation of the identity matrix whose parameters are the entries of X, then the deformation $A(y_1, \dots, y_t)$ is called *versal*. A versal deformation $A(y_1, \dots, y_t)$ with the minimum number t is called *miniversal*.

For example, all matrices $J_3(\lambda) + X$ that are close to $J_3(\lambda)$ can be reduced to the form

$$\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ y_1 & y_2 & y_3 \end{bmatrix} \tag{18}$$

by similarity transformations that are close to the identity and depend analytically on the entries of X. The matrix (18) is a linear miniversal deformation of $J_3(\lambda)$.

Let us formulate Arnold's theorem. We denote by $0^{\uparrow}_{pq}$ (respectively, $0^{\downarrow}_{pq}$, $0^{\leftarrow}_{pq}$, and $0^{\rightarrow}_{pq}$) the $p \times q$ matrix, in which all entries are zero except for the entries of the first row (respectively, last row, first column, and last column) that are stars. We usually omit the indices p and q. For example, the second summand in (18) is $0^{\downarrow}_{33}$.

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

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

Define the matrix with stars:

$$\widetilde{J}_{k_1,\dots,k_s}(\lambda) := \begin{bmatrix} J_{k_1}(\lambda) + 0^{\leftarrow} & 0^{\leftarrow} & \dots & 0^{\leftarrow} \\ 0^{\downarrow} & J_{k_2}(\lambda) + 0^{\leftarrow} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0^{\leftarrow} \\ 0^{\downarrow} & \dots & 0^{\downarrow} & J_{k_s}(\lambda) + 0^{\leftarrow} \end{bmatrix}. \tag{19}$$

Arnold's linear miniversal deformations of Jordan matrices are given in the following theorem, which is proved in [2, Theorem 4.4]; see also [3, Section 3.3] and [4, §30].

Theorem 3.1 *(Arnold [2]). Let a Jordan matrix be written in the form*

$$J = J_{k_{11},\dots,k_{1s_1}}(\lambda_1) \oplus \dots \oplus J_{k_{l1},\dots,k_{ls_l}}(\lambda_l), \qquad \begin{matrix} k_{i1} \leqslant k_{i2} \leqslant \dots \leqslant k_{is_i}, \\ \lambda_1, \dots, \lambda_l \in \mathbb{C} \text{ are distinct.} \end{matrix}$$

Then all matrices $J + X$ that are sufficiently close to J can be simultaneously reduced by some similarity transformation

$$J + X \mapsto S(X)^{-1}(J + X)S(X), \qquad \begin{matrix} S(X) \text{ is analytic} \\ \text{at } 0 \text{ and } S(0) = I, \end{matrix} \tag{20}$$

to the form

$$\widetilde{J} := \widetilde{J}_{k_{11},\dots,k_{1s_1}}(\lambda_1) \oplus \dots \oplus \widetilde{J}_{k_{l1},\dots,k_{ls_l}}(\lambda_l), \tag{21}$$

in which the stars are replaced by complex numbers that depend analytically on the entries of X at 0. The number of stars is minimum that can be achieved by similarity transformations of the form (20); this number is equal to the codimension of the similarity class of J.

A constructive proof of Theorem 3.1 by elementary transformations is given by Klimenko and Sergeichuk [31]. Many applications of miniversal deformations are given by Mailybaev [26,32–34]; he constructs a smooth similarity transformation (20) in the form of Taylor series. The radius of a neighborhood of J in which all matrices $J + X$ are reduced to the form (21) by transformations (20) is calculated in [11].

3.2. Miniversal deformations of matrix pairs

Denote by Z_{pq} the $p \times q$ matrix with $p \leqslant q$, in which the first $q-p$ entries of the first row are the stars and the other entries are zeros:

$$Z_{pq} := \begin{bmatrix} * & \dots & * & 0 & \dots & 0 \\ & & & & \ddots & \vdots \\ & 0 & & & & 0 \end{bmatrix}, \qquad p \leqslant q;$$

we usually omit the indices p and q.

Arnold's notion of miniversal deformations of matrices under similarity is naturally extended to matrix pairs under equivalence. A linear miniversal deformation of complex matrix pencils was first constructed by Edelman, Elmroth, and Kågström in the article [21], which was awarded the SIAM Linear Algebra Prize 2000 for the most outstanding paper published in 1997–1999. Their miniversal deformations contain repeating parameters (see (3)), which complicates their use in the proof of Theorem I. We use the following miniversal deformations.

Theorem 3.2 *(García-Planas, Sergeichuk [27, Theorem 4.1]). Let $\mathcal{A}$ be a Kronecker pair of form (16), in which $\lambda_1, \dots, \lambda_{t-1} \in \mathbb{C}$ are distinct and $\lambda_t = \infty$. Then all matrix pairs $\mathcal{A} + \mathcal{X}$ that are sufficiently close to $\mathcal{A}$ can be simultaneously reduced by some equivalence transformation*

$$\mathcal{A} + \mathcal{X} \mapsto R(\mathcal{X})^{-1}(\mathcal{A} + \mathcal{X})S(\mathcal{X}), \qquad \begin{gathered} \textit{matrices } R(\mathcal{X}) \textit{ and } S(\mathcal{X}) \\ \textit{are analytic at } (0,0), \\ R(0,0) = I, \textit{ and } S(0,0) = I, \end{gathered} \tag{22}$$

to the form

$$\left(\left[\begin{array}{cccc|c|c|ccc} L^T_{m_1} & & & 0 & & 0^{\downarrow} & & & \\ & L^T_{m_2} & & & & 0^{\downarrow} & & & \\ & & \ddots & & 0 & \vdots & 0^{\rightarrow} & \dots & 0^{\rightarrow} \\ & & & L^T_{m_{\underline{s}}} & & 0^{\downarrow} & & & \\ \hline & & & & I & 0 & & 0 & \\ \hline & & & & & \widetilde{J}_0 & 0^{\rightarrow} & \dots & 0^{\rightarrow} \\ \hline & & & & & & L_{n_{\overline{s}}} & & 0 \\ & & 0 & & & & & \ddots & \\ & & & & & & & L_{n_2} & \\ & & & & & & & & L_{n_1} \end{array}\right], \left[\begin{array}{cccc|c|c|ccc} R^T_{m_1} & Z & \dots & Z & 0^{\uparrow} & & & 0^{\uparrow} & \\ & R^T_{m_2} & \ddots & \vdots & 0^{\uparrow} & 0 & & 0^{\uparrow} & \\ & & \ddots & Z & \vdots & & & \vdots & \\ & & & R^T_{m_{\underline{s}}} & 0^{\uparrow} & & & 0^{\uparrow} & \\ \hline & & & & \widetilde{J} & 0 & 0^{\leftarrow} & \dots & 0^{\leftarrow} \\ \hline & & & & & I & & 0 & \\ \hline & & & & & & R_{n_{\overline{s}}} & Z^T \ \dots & Z^T \\ & & 0 & & & & & \ddots \quad \ddots & \vdots \\ & & & & & & & R_{n_2} & Z^T \\ & & & & & & & & R_{n_1} \end{array}\right] \right) \tag{23}$$

in which

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

Please cite this article in press as: V. Futorny et al., Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

(see (19)) and the stars are replaced by complex numbers that depend analytically on the entries of the pair $\mathcal{X}$ at $(0,0)$. The number of stars is minimum that can be achieved by equivalence transformations of the form (22).

Thus, the family of matrix pairs (23) is a linear miniversal deformation of the Kronecker matrix pair (16). The summands of each of the types in (16) and the corresponding horizontal and vertical strips in (23) can be absent.

By a *miniversal pair* we mean a matrix pair that is obtained from (23) by replacing its stars by complex numbers. We use the Frobenius matrix norm

$$\|[a_{ij}]\| := \sqrt{\sum\nolimits_{ij} |a_{ij}|^2}, \qquad a_{ij} \in \mathbb{C}. \tag{24}$$

For a matrix pair $\mathcal{A} = (A, A')$, we write $\|\mathcal{A}\| := \sqrt{\|A\|^2 + \|A'\|^2}$ and define its neighborhood

$$N_r(\mathcal{A}) := \{\mathcal{B} \mid \|\mathcal{B} - \mathcal{A}\| < r\},$$

in which r is a positive real number.

Remark 3.1. Let $\mathcal{A}$ be the matrix pair from Theorem 3.2. Let $N_r(\mathcal{A})$ be its neighborhood, in which all pairs are reduced to the form (23) by an analytic transformation $\mathcal{A} + \mathcal{X} \mapsto \mathcal{A} + \widehat{\mathcal{X}}$ from (22). Since it is analytic, there is a positive $c \in \mathbb{R}$ such that

$$\|\widehat{\mathcal{X}}\| \leqslant c\|\mathcal{X}\| \qquad \text{for all } \mathcal{A} + \mathcal{X} \in N_r(\mathcal{A}).$$

Hence, each pair in $N_r(\mathcal{A})$ is equivalent to a miniversal pair from $N_{cr}(\mathcal{A})$. Thus, if a Kronecker pair $\mathcal{B}$ is equivalent to a pair in an arbitrarily small neighborhood of $\mathcal{A}$, then $\mathcal{B}$ is equivalent to a *miniversal* pair in an arbitrarily small neighborhood of $\mathcal{A}$. We use this fact in the proof of Theorem I.

Miniversal deformations were also constructed for matrices under congruence [17] and *congruence [19], for pairs of symmetric matrices under congruence [15], for pairs of skew-symmetric matrices under congruence [14], and for matrix pairs under contragredient equivalence [27].

3.3. Weyr canonical form

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.

Let A be a matrix with the single eigenvalue 0, and let $(m_1, m_2, \dots)$ be its Weyr characteristic for 0. In the proof of Theorem 6.1, we use the fact that A is similar to the matrix

Please cite this article in press as: V. Futorny et al., Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

$$W=\begin{bmatrix}0_{m_1} & F_1 & & 0\\ & 0_{m_2} & \ddots & \\ & & \ddots & F_{k-1}\\ 0 & & & 0_{m_k}\end{bmatrix},\qquad F_i:=\begin{bmatrix}I_{m_{i+1}}\\0\end{bmatrix},\tag{25}$$

which is permutation similar to the Jordan canonical form of A. Sergeichuk [39] suggested to call W the *Weyr canonical form* of A. Now this term is generally accepted; see historical remarks in [36, pp. 80–82]. Each matrix with distinct eigenvalues $\lambda_1,\dots,\lambda_s$ is similar to a Weyr matrix $(\lambda_1 I+W_1)\oplus\cdots\oplus(\lambda_s I+W_s)$, in which $W_1,\dots,W_s$ are of the form (25).

Applying the permutation of rows and columns of (19) that transform $J_{k_1,\dots,k_s}(\lambda)$ to its Weir canonical form, Klimenko and Sergeichuk [30] obtain a matrix in which all stars are on the main diagonal and under it. Thus, if we replace each Jordan matrix $\mathcal{D}_{k_{i1}}(\lambda_i)\oplus\cdots\oplus\mathcal{D}_{k_{is_i}}(\lambda_i)$ in (16) by the transpose of its Weyr canonical form, and use such matrices instead of J and J_0 in (23), then we obtain a matrix in which all stars are on the main diagonal and over it. The obtained matrix defines an upper triangular miniversal deformation of a Kronecker pair in which Weyr matrices are used instead of Jordan matrices.

4. A direct proof of a weak form of Theorem IV

Due to the following theorem, which is a weak form of Theorem IV, it suffices to prove Theorem I for all matrix pairs (16) 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 4.1. *Let $\mathcal{A}=([A_{ij}],[A'_{ij}])$ be a Kronecker pair of form (16) partitioned such that the pairs of diagonal blocks $(A_{11},A'_{11}),(A_{22},A'_{22}),\dots$ are the direct summands (17). 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 obtained from $\mathcal{A}$ by an arbitrarily small perturbation of only one pair (A_{ij},A'_{ij}) with $i<j$ that is not contained in $\mathcal{D}_1,\dots,\mathcal{D}_t$, or of only one pair (A_{ij},A'_{ij}) from $\mathcal{D}_1,\dots,\mathcal{D}_t$.

Proof. Besides the partition of the matrices of $\mathcal{A}=(A,A')$ into the blocks A_{ij} and A'_{ij}, we also consider the partition of A and A' into the *superblocks* obtained by joining all strips that correspond to the same eigenvalue. Thus, the diagonal superblocks form the matrix pairs

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

Please cite this article in press as: V. Futorny et al., Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

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

$$\mathcal{B}_1 = (B_1, B_1'),\ \mathcal{B}_2 = (B_2, B_2'),\ \dots \tag{26}$$

of pairs from $\langle\mathcal{B}\rangle$ that converges to $\mathcal{A}$. All matrix pairs close enough to $\mathcal{A}$ are reduced to the miniversal form (23) by a smooth equivalence transformation that preserves $\mathcal{A}$. Hence, all pairs (26) can be taken in the miniversal form (23), which is *upper superblock triangular.*

We say that a block (superblock) of B_i or B_i' in (26) is *perturbed* if it differs from the corresponding block (superblock) of A or A'.

Case 1: There are infinite many pairs (26)*, in which at least one upper diagonal superblock is perturbed.*

Then there is a partition

$$\mathcal{A} = \left(\begin{bmatrix} M & O \\ 0 & N \end{bmatrix}, \begin{bmatrix} M' & O' \\ 0 & N' \end{bmatrix}\right) \quad \begin{matrix} M \text{ and } M' \text{ are } m \times m', \\ O \text{ and } O' \text{ are zero} \end{matrix} \tag{27}$$

that is coarser than the partition into superblocks, with the property: O or O' is perturbed infinitely many times in the sequence (26). We can suppose that O or O' is perturbed in *each* pair (26).

Let the partition

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

be conformal with (27). Write $\xi_i := (\|O_i\| + \|O_i'\|)^{-1}$, in which $\|\cdot\|$ is the Frobenius matrix norm (24). Define the equivalent pair

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

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

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

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

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

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

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 and M_1', and $c\times d$ is the size of N_3 and 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 $\mathcal{Y}$ is a desired pair, it suffices to prove that $\langle\mathcal{Y}\rangle\neq\langle\mathcal{A}\rangle$ (which implies $\langle\mathcal{B}\rangle=\langle\mathcal{Y}\rangle>\langle\mathcal{A}\rangle$ because $\langle\mathcal{B}\rangle$ is an immediate successor of $\langle\mathcal{A}\rangle$).

On the contrary, suppose that $\langle\mathcal{Y}\rangle=\langle\mathcal{A}\rangle$. Since $\mathcal{Y}_\varepsilon$ is equivalent to $\mathcal{Y}$, we have $\mathcal{Y}_\varepsilon\in\langle\mathcal{A}\rangle$ for each ε. Hence there exist nonsingular matrices, which we take in the form $I+R_\varepsilon$ and $I+S_\varepsilon$, such that

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

By Lipschitz's property for matrix pairs (see [38] or [1]), we can choose the matrices R_ε and S_ε and a positive constant $c\in\mathbb{R}$ such that

$$\|R_\varepsilon\|<\varepsilon c\quad\text{and}\quad\|S_\varepsilon\|<\varepsilon c\qquad\text{for all }\varepsilon,\tag{28}$$

in which $\|\cdot\|$ is the Frobenius matrix norm (24).

The pair $\mathcal{Y}_\varepsilon$ is in the miniversal form (23) for (27) since all nonzero entries of Q and Q' are at the places of some stars. By the construction of miniversal deformation (23), which is given in [27, Theorem 4.1], the pair

$$\Delta\mathcal{Y}_\varepsilon := \mathcal{Y}_\varepsilon - \mathcal{A} = \varepsilon\left(\left[\begin{array}{c|c} 0 & \begin{smallmatrix} 0&0&0\\0&X&0\\0&0&0\end{smallmatrix} \\ \hline 0 & 0 \end{array}\right], \left[\begin{array}{c|c} 0 & \begin{smallmatrix} 0&0&0\\0&X'&0\\0&0&0\end{smallmatrix} \\ \hline 0 & 0 \end{array}\right]\right) = R_\varepsilon\mathcal{A} + \mathcal{A}S_\varepsilon + R_\varepsilon\mathcal{A}S_\varepsilon \tag{29}$$

does not belong to the space

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

(which is the tangent space to $\langle\mathcal{A}\rangle$ at $\mathcal{A}$). Thus,

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

(which is the distance from $\mathcal{Y}_\varepsilon$ to the affine space $\{\mathcal{A} + R\mathcal{A} + \mathcal{A}S \mid R, S\}$).

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 (29), $\Delta\mathcal{Y}_\varepsilon = \varepsilon\Delta\mathcal{Y}_1$, and so $\varepsilon d_1 = \|\Delta\mathcal{Y}_\varepsilon - (\varepsilon R')\mathcal{A} - \mathcal{A}(\varepsilon S')\| = d_\varepsilon$. By (28),

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

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 (26) in which at least one upper diagonal superblock is perturbed.

Let $\mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \dots$ be the pairs of diagonal superblocks of $\mathcal{A}$, then $\mathcal{A} = \mathcal{A}^{(1)} \oplus \mathcal{A}^{(2)} \oplus \cdots$. We can suppose that all upper diagonal superblocks are not perturbed, and so $\mathcal{B}_i := \mathcal{B}_i^{(1)} \oplus \mathcal{B}_i^{(2)} \oplus \cdots$, in which $\mathcal{B}_i^{(1)}, \mathcal{B}_i^{(2)}, \dots$ are the pairs of perturbed diagonal superblocks of $\mathcal{B}_i$ in (26).

Since all $\mathcal{B}_i \sim \mathcal{B}$ (the symbol $\sim$ means "equivalent"), we can suppose that $\mathcal{B}_1^{(l)} \sim \mathcal{B}_2^{(l)} \sim \cdots$ for each l. Since $\mathcal{A} \nsim \mathcal{B}$, we have $\mathcal{A}^{(l)} \nsim \mathcal{B}_1^{(l)} \sim \mathcal{B}_2^{(l)} \sim \cdots$ 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 $\langle\mathcal{C}_i\rangle = \langle\mathcal{C}_1\rangle > \langle\mathcal{A}\rangle$. Moreover, $\langle\mathcal{B}\rangle \geqslant \langle\mathcal{C}_1\rangle$ because

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

There is no intermediate orbit between $\langle\mathcal{A}\rangle$ and $\langle\mathcal{B}\rangle$, and so $\langle\mathcal{B}\rangle = \langle\mathcal{C}_1\rangle$. $\square$

5. Perturbations of direct sums of two indecomposable Kronecker pairs

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

Theorem 5.1. *For the Kronecker pair*

$$\mathcal{L}_m^T \oplus \mathcal{L}_n^T, \qquad m \leqslant n, \tag{30}$$

its miniversal deformation from Theorem 3.2 is given by the matrix pair

$$\left(\left[\begin{array}{c|c} L_m^T & 0 \\ \hline 0 & L_n^T \end{array}\right], \left[\begin{array}{c|c} R_m^T & \begin{matrix} \alpha_1 \ \dots \ \alpha_{n-1} \\ 0 \end{matrix} \\ \hline 0 & R_n^T \end{array}\right] \right), \tag{31}$$

in which

$$(\alpha_1, \dots, \alpha_{n-1}) = (*, \dots, *, \underbrace{0, \dots, 0}_{m}). \tag{32}$$

(a) *If the stars in (32) are complex numbers that are not all zero, then (31) is equivalent to one of the pairs*

$$\mathcal{L}_{m+r}^T \oplus \mathcal{L}_{n-r}^T, \qquad m + r \leqslant n - r,\ r \geqslant 1. \tag{33}$$

(b) *Each pair (33) is equivalent to a pair of the form*

$$\left(\begin{bmatrix} L_m^T & 0 \\ 0 & L_n^T \end{bmatrix}, \begin{bmatrix} R_m^T & \Delta_r(\varepsilon) \\ 0 & R_n^T \end{bmatrix} \right), \tag{34}$$

in which $\Delta_r(\varepsilon)$ is defined in (9) and ε is an arbitrary nonzero complex number.

(c) *The set of Kronecker canonical forms of all pairs in a sufficiently small neighborhood of (30) consists of the pairs*

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

Lemma 5.1. *Each pair of $n \times (n-1)$ matrices of the form*

$$\left(\begin{bmatrix} 1 & * & & * \\ 0 & 1 & \ddots & \\ & 0 & \ddots & * \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix}, \begin{bmatrix} * & * & & * \\ 1 & * & \ddots & \\ & 1 & \ddots & * \\ & & \ddots & * \\ 0 & & & 1 \end{bmatrix} \right) \tag{35}$$

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 (35) obtained by removing the last row and last column in the matrices of the pair (35). Reasoning by induction on n, we suppose that the subpair $\mathcal{P}$ is reduced to $\mathcal{L}_{n-1}^T$ by simultaneous additions of columns of its matrices from left to right and simultaneous additions of rows from the bottom to up. We obtain (35) in which all entries that are marked by stars are zero except for some entries of the last columns. We make zero the entries 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. □

Proof of Theorem 5.1. (a) Let the stars in (32) be complex numbers that are not all zero, and let α_s be the first nonzero entry. Then

$$1 \leqslant s < n - m. \tag{36}$$

Let (C, D) be the matrix (31); we will reduce it by simultaneous elementary transformations to the form (33). We usually specify only transformations with one of the matrices C and D; it is understood 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 (32) with $\alpha_s = 1$. Consider the obtained pair

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

$$=\left(\begin{array}{c|c:c:c}
{\scriptstyle 1\ \ m-1} & {\scriptstyle 1\ \ \ s} & {\scriptstyle s+m-1} & \\
\hline
\begin{smallmatrix}1&&\\0&\ddots&\\&\ddots&1\\&&0\end{smallmatrix} & & & \\
\hline
 & \begin{smallmatrix}1&&&\\0&\ddots&&\\&\ddots&1&\\&&0&1\end{smallmatrix} & & \\
\hdashline
\begin{smallmatrix}\emptyset&&\\0&\ddots&\\&\ddots&\emptyset\\&&0\end{smallmatrix} & \begin{smallmatrix}0\end{smallmatrix} & \boxed{\begin{smallmatrix}1&&\\0&\ddots&\\&\ddots&1\\&&0\end{smallmatrix}} & \begin{smallmatrix}1\end{smallmatrix} \\
\hdashline
 & & & \begin{smallmatrix}0&&\\ \ddots&\ddots&\\&&1\\&&0\end{smallmatrix}
\end{array}\ ,\
\begin{array}{c|c:c:c|l}
{\scriptstyle 1\ \ m-1} & {\scriptstyle 1\ \ \ s} & {\scriptstyle s+m-1} & & \\
\hline
\begin{smallmatrix}0&&\\1&\ddots&\\&\ddots&0\\&&1\end{smallmatrix} & 0\ \dots\ 0\ 1 & *\ \dots\ * & *\ \dots\ * & {\scriptstyle 1}\\
 & & & & {\scriptstyle m}\\
\hline
 & \begin{smallmatrix}0&&&\\1&\ddots&&\\&\ddots&0&\\&&1&0\end{smallmatrix} & & & {\scriptstyle 1}\\
 & & & & {\scriptstyle s}\\
\hdashline
\begin{smallmatrix}0&&\\\emptyset&\ddots&\\&\ddots&0\\&&\emptyset\end{smallmatrix} & \begin{smallmatrix}1\end{smallmatrix} & \boxed{\begin{smallmatrix}0&&\\1&\ddots&\\&\ddots&0\\&&1\end{smallmatrix}} & \begin{smallmatrix}0\end{smallmatrix} & {\scriptstyle s+1}\\
 & & & & {\scriptstyle s+m}\\
\hdashline
 & & & \begin{smallmatrix}1&&\\ \ddots&\ddots&\\&&0\\&&1\end{smallmatrix} &
\end{array}\right) \tag{37}$$

We denote by $\emptyset$ the entries in (37) that are transformed to -1 and then are restored to 0 during the following simultaneous elementary transformations, which make zero the entry "1" under α_s:

- 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 in the block $(2,1)$ (see (37)).
- 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

$$(C,D) = \left(\left[\begin{array}{c|c|c} \begin{matrix}1&&\\0&\ddots&\\&\ddots&1\\&&0\end{matrix} & & \\ \hline & \begin{matrix}1&&&\\&\ddots&&\\&&1&\\&&&1\end{matrix} & \\ \hline & & \begin{matrix}1&&\\0&\ddots&\\&\ddots&1\\&&0\end{matrix} \end{array}\right], \left[\begin{array}{c|c|c} \begin{matrix}*&\dots&*\\1&&0\\&\ddots&\\&&1\end{matrix} & \begin{matrix}0&\dots&0&1\end{matrix} & \begin{matrix}*&\dots&*\end{matrix} \\ \hline & \begin{matrix}0&&&\\1&\ddots&&\\&\ddots&0&\\&&1&0\end{matrix} & \\ \hline \begin{matrix}*&\dots&*\end{matrix} & & \begin{matrix}*&\dots&*\\1&&0\\&\ddots&\\&&1\end{matrix} \end{array}\right] \right), \tag{38}$$

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

$$\begin{aligned} (C,D) &= \left(\left[\begin{array}{c|c|c} C_{11} & C_{12} & C_{13} \\ \hline C_{21} & C_{22} & C_{23} \\ \hline C_{31} & C_{32} & C_{33} \end{array}\right], \left[\begin{array}{c|c|c} D_{11} & D_{12} & D_{13} \\ \hline D_{21} & D_{22} & D_{23} \\ \hline D_{31} & D_{32} & D_{33} \end{array}\right] \right) \\ &= \left(\left[\begin{array}{c|c|c} \begin{matrix}1&&&\\0&\ddots&&\\&\ddots&1&\\&&0&1\end{matrix} & & \\ \hline \begin{matrix}&&&0\end{matrix} & \begin{matrix}1&&&\\0&1&&\\&0&\ddots&\\&&\ddots&1\\&&&0\end{matrix} & \\ \hline & \begin{matrix}0&*&\dots&*\\&0&\ddots&\vdots\\&&\ddots&*\\&&&0\end{matrix} & \begin{matrix}1&&\\0&\ddots&\\&\ddots&1\\&&0\end{matrix} \end{array}\right], \left[\begin{array}{c|c|c} \begin{matrix}0&&&\\1&\ddots&&\\&\ddots&0&\\&&1&0\end{matrix} & & \\ \hline \begin{matrix}&&&1\end{matrix} & \begin{matrix}*&*&\dots&*\\1&&&0\\&1&&\\&&\ddots&\\&&&1\end{matrix} & \begin{matrix}*&\dots&*\end{matrix} \\ \hline & \begin{matrix}*&*&\dots&*\\&*&\ddots&\vdots\\&&\ddots&*\\&&&*\end{matrix} & \begin{matrix}*&\dots&*\\1&&0\\&\ddots&\\&&1\end{matrix} \end{array}\right] \right), \end{aligned} \tag{39}$$

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

Using transformations from Lemma 5.1, 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 (39) in which all stars in horizontal strips 3 of C and D are zero.

Using Lemma 5.1, 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 (31) and that is obtained from $\mathcal{L}^T_{m+s} \oplus \mathcal{L}^T_{n-s}$ by an arbitrarily small perturbation, in which s as in (37) and satisfies (36). If $m+s > n-s$, then we interchange $\mathcal{L}^T_{m+s}$ and $\mathcal{L}^T_{n-s}$ and reduce the obtained pair by equivalence transformations to its miniversal form.

We obtain

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

in which the stars are sufficiently small complex numbers. By (36),

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

We repeat this procedure until we obtain a pair

$$\left(\left[\begin{array}{c|c} L^T_{m^{(l)}} & 0 \\ \hline 0 & L^T_{n^{(l)}} \end{array}\right], \left[\begin{array}{c|c} R^T_{m^{(l)}} & \begin{matrix} * \cdots * \\ 0 \end{matrix} \\ \hline 0 & R^T_{n^{(l)}} \end{array}\right]\right) \tag{40}$$

in which all stars are zero, and $m < m^{(l)} \leqslant n^{(l)}$. Thus, (40) is of the form (33).

(b) Let $\mathcal{L}^T_{m+r} \oplus \mathcal{L}^T_{n-r}$ be the pair (33); we must prove that it is equivalent to (34). We divide the first horizontal strips of (34) by ε, then multiply the first vertical strips by ε, and obtain the pair (37) in which all stairs are zero. The obtained pair is reduced as above to (38) in which all stairs are zero. This pair is permutation equivalent to $\mathcal{L}^T_{m+r} \oplus \mathcal{L}^T_{n-r}$.

(c) This statement follows from (a), (b), Theorem 3.2, and Remark 3.1. □

Example 5.1. If $m = 2$ and $n = 8$ in (31), then $(\alpha_1, \dots, \alpha_7) = (*, *, *, *, *, 0, 0)$. The pairs

$$\left(\begin{bmatrix} L^T_2 & 0 \\ 0 & L^T_8 \end{bmatrix}, \begin{bmatrix} R^T_2 & \Delta_r(\varepsilon) \\ 0 & R^T_8 \end{bmatrix}\right) \quad \text{with } \varepsilon \neq 0 \text{ and } r = 1,\, 2,\, 3,\, 4,\, 5$$

are equivalent to $\mathcal{L}^T_3 \oplus \mathcal{L}^T_7$, $\mathcal{L}^T_4 \oplus \mathcal{L}^T_6$, $\mathcal{L}^T_5 \oplus \mathcal{L}^T_5$, $\mathcal{L}^T_6 \oplus \mathcal{L}^T_4$, $\mathcal{L}^T_7 \oplus \mathcal{L}^T_3$, respectively. We obtain both $\mathcal{L}^T_3 \oplus \mathcal{L}^T_7$ and $\mathcal{L}^T_7 \oplus \mathcal{L}^T_3$ since they are not reduced one to the other by equivalence transformations that are close to the identity; whereas the pairs that are close to $\mathcal{A}$ are reduced to the form (23) by smooth transformations (22) that are close to the identity.

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

Theorem 5.2. *For the Kronecker pair*

$$\mathcal{L}_m \oplus \mathcal{L}_n, \qquad m \leqslant n, \tag{41}$$

its miniversal deformation from Theorem 3.2 is given by the matrix pair (C^T, D^T), in which (C, D) is the pair (31).

(a) *If the stars in (32) are complex numbers that are not all zero, then (C^T, D^T) is equivalent to one of the pairs*

$$\mathcal{L}_{m+r} \oplus \mathcal{L}_{n-r}, \qquad m + r \leqslant n - r,\ r \geqslant 1. \tag{42}$$

(b) *Each pair (42) is equivalent to a pair of the form*

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

in which $\Delta_r(\varepsilon)$ is defined in (9) and ε is an arbitrary nonzero complex number.

(c) *The set of Kronecker canonical forms of all pairs in a sufficiently small neighborhood of (41) consists of the pairs*

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

Proof. This theorem is obtained from Theorem 5.1 by matrix transposition. □

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

Theorem 5.3. *For the Kronecker pair*

$$\mathcal{L}_m^T \oplus \mathcal{D}_n(\lambda), \qquad \lambda \in \mathbb{C} \cup \infty, \tag{43}$$

the matrix pair (23) of its miniversal deformation without stars in the diagonal blocks is

$$\mathcal{L}_m^T \oplus \mathcal{D}_n(\lambda) = \begin{cases} \left(\left[\begin{array}{c|c} L_m^T & 0 \\ \hline 0 & I_n \end{array} \right], \left[\begin{array}{c|c} R_m^T & \begin{matrix} \alpha_1 \dots \alpha_n \\ 0 \end{matrix} \\ \hline 0 & J_n(\lambda) \end{array} \right] \right) & \text{if } \lambda \in \mathbb{C} \\ \left(\left[\begin{array}{c|c} L_m^T & \begin{matrix} 0 \\ \alpha_1 \dots \alpha_n \end{matrix} \\ \hline 0 & J_n(0) \end{array} \right], \left[\begin{array}{c|c} R_m^T & 0 \\ \hline 0 & I_n \end{array} \right] \right) & \text{if } \lambda = \infty \end{cases} \tag{44}$$

in which $(\alpha_1, \dots, \alpha_n) = (*, \dots, *)$.

(a) *Let the stars in* (44) *be complex numbers that are not all zero, and let* α_s *be the first nonzero element in* $(\alpha_1, \dots, \alpha_n)$. *Then* (44) *is equivalent to the pair*

$$\mathcal{L}^T_{m+n-s+1} \oplus \mathcal{D}_{s-1}(\lambda). \tag{45}$$

(b) *Each pair* (45) *with* $s \in \{1, \dots, n\}$ *is equivalent to a pair of the form*

$$\begin{array}{ll} \left(\begin{bmatrix} L_m^T & 0 \\ 0 & I_n \end{bmatrix}, \begin{bmatrix} R_m^T & \Delta_s(\varepsilon) \\ 0 & J_n(\lambda) \end{bmatrix} \right) & \text{if } \lambda \in \mathbb{C} \\ \left(\begin{bmatrix} L_m^T & \nabla_s(\varepsilon) \\ 0 & J_n(0) \end{bmatrix}, \begin{bmatrix} R_m^T & 0 \\ 0 & I_n \end{bmatrix} \right) & \text{if } \lambda = \infty \end{array} \tag{46}$$

in which $\Delta_r(\varepsilon)$ *and* $\nabla_r(\varepsilon)$ *are defined in* (9) *and* ε *is an arbitrary nonzero complex number.*

(c) *The set of Kronecker canonical forms of all pairs obtained by perturbations of the blocks* $(1, 2)$ *in* (43) *consists of the pairs*

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

Proof. Let (A, B) be the pair (43) with $\lambda = \infty$. Write

$$Z_p := \begin{bmatrix} 0 & & 1 \\ & .\cdot^{\cdot} & \\ 1 & & 0 \end{bmatrix} \quad (p\text{-by-}p).$$

Since $(R_m^T, L_m^T) = Z_m(L_m^T, R_m^T)Z_{m-1}$, we have that (B, A) is equivalent to the pair (43) with $\lambda = 0$. Therefore, it suffices to prove the theorem for $\lambda \in \mathbb{C}$.

Let $(A, B(\lambda))$ be the pair (43) with $\lambda \in \mathbb{C}$. By Lemma 5.1, 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 $\lambda = 0$. Then (44) in which $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ is

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

$$= \left(\left[\begin{array}{cccc|cccc} 1 & & & & & & & \\ 0 & \ddots & & & & & & \\ & \ddots & 1 & & & 0 & & \\ & & 0 & & & & & \\ \hline & & & & 1 & & & \\ & & & & & 1 & & \\ & 0 & & & & & \ddots & \\ & & & & & & & 1 \end{array} \right], \left[\begin{array}{cccc|cccc} 0 & & & & \alpha_1 & \alpha_2 & \dots & \alpha_n \\ 1 & \ddots & & & & & & \\ & \ddots & 0 & & & 0 & & \\ & & 1 & & & & & \\ \hline & & & & 0 & 1 & & \\ & & & & & 0 & \ddots & \\ & 0 & & & & & \ddots & 1 \\ & & & & & & & 0 \end{array} \right] \right). \tag{47}$$

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

$$K_n := \begin{bmatrix} \kappa_1 & \kappa_2 & \ddots & \kappa_n \\ & \kappa_1 & \ddots & \ddots \\ & & \ddots & \kappa_2 \\ 0 & & & \kappa_1 \end{bmatrix}, \qquad \kappa_1, \dots, \kappa_n \in \mathbb{C}. \tag{48}$$

The equivalence transformation

$$(I_m \oplus K_n^{-1})(C, D)(I_{m-1} \oplus K_n), \qquad \kappa_1 \neq 0 \tag{49}$$

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

$$(\alpha_1, \dots, \alpha_n)K_n = (\alpha_1\kappa_1,\ \alpha_1\kappa_2 + \alpha_2\kappa_1,\ \dots,\ \alpha_1\kappa_n + \dots + \alpha_n\kappa_1) \tag{50}$$

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 (50), we make $(\alpha_1, \dots, \alpha_n) = (0, \dots, 0, 1, 0, \dots, 0)$ with "1" at the position s.

Let first $s = 1$. Then

$$(Z_m \oplus I_n)(C, D)(Z_{m-1} \oplus I_n) = (R_{m+n}^T, L_{m+n}^T) \sim (L_{m+n}^T, R_{m+n}^T),$$

which is a pair of the form (45).

Let now $s \geqslant 2$. The "1" under $\alpha_s = 1$ is the $(s-1, s)$th entry of the block D_{22} (see (47)). We make zero this entry of D_{22} by the following elementary transformations:

- *Case 1: $m < s$.* 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

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

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

Please cite this article in press as: V. Futorny et al., Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

- *Case 2: $m \geqslant s$.* 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

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

Therefore, for each s the pair (C, D) is reduced to the pair (C', D') that is obtained from (47) by replacing $(\alpha_1, \dots, \alpha_n)$ by $(0, \dots, 0, 1, 0, \dots, 0)$ with $\alpha_s = 1$ and by replacing the entry "1" under α_s by 0. Then $(Z_m \oplus I_n)(C', D')(Z_{m-1} \oplus I_n) = (R^T_{m+n-s+1}, L^T_{m+n-s+1}) \oplus (I_{s-1}, J_{s-1}(0)) \sim \mathcal{L}^T_{m+n} \oplus \mathcal{D}_{s-1}(0)$, which is a pair of the form (45).

(b) The pair (46) with $\lambda = 0$ is the pair (47) in which $(\alpha_1, \dots, \alpha_n) = (0, \dots, 0, \varepsilon, 0, \dots, 0)$ with $\varepsilon \neq 0$ at the place s. Reasoning as in part (a), we reduce it to the pair (45).

(c) Because of the statement (a), it is sufficient to prove that all pairs in a sufficiently small neighborhood of $\mathcal{L}^T_m \oplus \mathcal{D}_n(0)$ that are obtained by perturbations of its blocks (1,2) are reduced to the form (44) with $\lambda = 0$ by transformations (22). To keep matters clear, let us prove it for $m = 3$ and $n = 2$; that is, for the pairs

$$\left(\left[\begin{array}{cc|cc} 1&0&x_{11}&x_{12} \\ 0&1&x_{21}&x_{22} \\ 0&0&x_{31}&x_{32} \\ \hline 0&0&1&0 \\ 0&0&0&1 \end{array}\right], \left[\begin{array}{cc|cc} 0&0&y_{11}&y_{12} \\ 1&0&y_{21}&y_{22} \\ 0&1&y_{31}&y_{32} \\ \hline 0&0&0&1 \\ 0&0&0&0 \end{array}\right] \right), \tag{51}$$

in which all x_{ij} and y_{ij} are sufficiently small complex numbers. We successively make $(x_{31}, x_{32}) = (0, 0)$ by adding rows of the second strip, $(y_{31}, y_{32}) = (0, 0)$ by adding the second column, $(x_{21}, x_{22}) = (0, 0)$ by adding rows of the second strip, $(y_{21}, y_{22}) = (0, 0)$ by adding the first column, and $(x_{11}, x_{12}) = (0, 0)$ by adding rows of the second strip. We obtain (51), in which all x_{ij} and y_{ij} are zero except to y_{11} and y_{12}. $\square$

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

Theorem 5.4. *For the Kronecker pair*

$$\mathcal{L}_m \oplus \mathcal{D}_n(\lambda), \qquad \lambda \in \mathbb{C} \cup \infty, \tag{52}$$

the matrix pair (23) of its miniversal deformation without stars in the diagonal blocks is

$$\mathcal{L}_m \oplus \mathcal{D}_n(\lambda) = \begin{cases} \left(\left[\begin{array}{c|c} L_m & 0 \\ \hline 0 & I_n \end{array}\right], \left[\begin{array}{cc|c} \multicolumn{2}{c|}{R_m} & 0 \\ \hline \begin{matrix}\alpha_1\\ \vdots \\ \alpha_n\end{matrix} & 0 & J_n(\lambda) \end{array}\right] \right) & \text{if } \lambda \in \mathbb{C} \\ \left(\left[\begin{array}{cc|c} \multicolumn{2}{c|}{L_m} & 0 \\ \hline \begin{matrix}\alpha_1\\ \vdots \\ \alpha_n\end{matrix} & 0 & J_n(0) \end{array}\right], \left[\begin{array}{c|c} R_m & 0 \\ \hline 0 & I_n \end{array}\right] \right) & \text{if } \lambda = \infty \end{cases} \tag{53}$$

in which $(\alpha_1, \dots, \alpha_n) = (*, \dots, *)$.

(a) *Let the stars in (53) be complex numbers that are not all zero, and let α_r be the last nonzero element in $(\alpha_1, \dots, \alpha_n)$. Then (44) is equivalent to the pair*

$$\mathcal{L}_{m+r} \oplus \mathcal{D}_{n-r}(\lambda). \tag{54}$$

(b) *Each pair (54) with $r \in \{1, \dots, n\}$ is equivalent to a pair of the form*

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

in which ε is an arbitrary nonzero complex number.

(c) *The set of Kronecker canonical forms of all pairs obtained by perturbations of the blocks $(1, 2)$ in (52) consists of the pairs*

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

Proof. The mapping

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

transforms the matrices from Theorem 5.3 to the matrices from Theorem 5.4. □

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

Theorem 5.5. *For the Kronecker pair*

$$\mathcal{D}_m(\lambda) \oplus \mathcal{D}_n(\lambda), \qquad m \leqslant n, \ \lambda \in \mathbb{C} \cup \infty, \tag{55}$$

the matrix pair (23) of its miniversal deformation with stars only in the blocks $(1, 2)$ is

$$\begin{cases} \left(I_{m+n}, \begin{bmatrix} J_m(\lambda) & A \\ 0 & J_n(\lambda) \end{bmatrix}\right) & \textit{if } \lambda \in \mathbb{C}, \\ \left(\begin{bmatrix} J_m(0) & A \\ 0 & J_n(0) \end{bmatrix}, I_{m+n}\right) & \textit{if } \lambda = \infty, \end{cases} \qquad A := \begin{bmatrix} \alpha_1 & & \\ \vdots & & 0 \\ \alpha_m & & \end{bmatrix}, \tag{56}$$

in which $\alpha_1, \dots, \alpha_m$ *are stars.*

(a) *Let* $\alpha_1, \dots, \alpha_m$ *in* (56) *be complex numbers that are not all zero. Let* α_r *be the last nonzero element in this sequence. Then* (56) *is equivalent to the pair*

$$\mathcal{D}_{m-r}(\lambda) \oplus \mathcal{D}_{n+r}(\lambda). \tag{57}$$

(b) *Each pair* (57) *with* $r \in \{1, \dots, m\}$ *is equivalent to the pair* (56)*, in which* $(\alpha_1, \dots, \alpha_m) = (0, \dots, 0, \varepsilon, 0, \dots, 0)$ *and* ε *is an arbitrary nonzero complex number in the* r*-th position.*

(c) *If a given Kronecker pair* $\mathcal{K}$ *is equivalent to a pair in an arbitrarily small neighborhood of* (55)*, then* $\mathcal{K}$ *has the form* $\mathcal{D}_{m-r}(\lambda) \oplus \mathcal{D}_{n+r}(\lambda)$ *with* $r \in \{0, \dots, m\}$.

Proof. This theorem follows from Theorem 6.2 by the reasons that are given at the beginning of Section 6. □

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

Theorem 5.6. *For the Kronecker pair*

$$\mathcal{L}_m^T \oplus \mathcal{L}_n, \tag{58}$$

the matrix pair (23) *of its miniversal deformation is*

$$\left(\left[\begin{array}{cccc|cccc} 1 & & & & & & & \alpha_1 \\ 0 & \ddots & & & & & & \alpha_2 \\ & \ddots & 1 & & & 0 & & \vdots \\ & & 0 & & & & & \alpha_m \\ \hline & & & & 1 & 0 & & \\ & 0 & & & & \ddots & \ddots & \\ & & & & & & 1 & 0 \end{array}\right], \left[\begin{array}{cccc|cccc} 0 & & & & \beta_1 & \beta_2 & \dots & \beta_n \\ 1 & \ddots & & & & & & \\ & \ddots & 0 & & & 0 & & \\ & & 1 & & & & & \\ \hline & & & & 0 & 1 & & \\ & 0 & & & & \ddots & \ddots & \\ & & & & & & 0 & 1 \end{array}\right]\right), \tag{59}$$

in which all α_i *and* β_j *are stars.*

(a) *Let the stars in* (59) *be complex numbers that are not all zero. Let*

$$\begin{gathered} (-\beta_1, \dots, -\beta_n, \alpha_1, \dots, \alpha_m) = \varepsilon(c_0, \dots, c_{r-1}, 1, \underbrace{0, \dots, 0}_{l \geqslant 0}), \quad \varepsilon \neq 0, \\ c_0 + c_1 x + \dots + c_{r-1} x^{r-1} + x^r = (x - \lambda_1)^{r_1} \cdots (x - \lambda_s)^{r_s} \end{gathered} \tag{60}$$

with distinct $\lambda_1,\dots,\lambda_s\in\mathbb{C}$*. Then* (59) *is equivalent to the pair*

$$\mathcal{D}_{r_1}(\lambda_1)\oplus\cdots\oplus\mathcal{D}_{r_s}(\lambda_s)\oplus\mathcal{D}_l(\infty). \tag{61}$$

(b) *Each pair* (61) *with distinct* $\lambda_1,\dots,\lambda_s\in\mathbb{C}$*, positive* $r_1,\dots,r_s$*, and* $l\geqslant 0$ *is equivalent to the pair* (59)*, in which* $\alpha_1,\dots,\alpha_m,\beta_1,\dots,\beta_n$ *are determined by* (60) *and an arbitrary nonzero* $\varepsilon\in\mathbb{C}$*.*

(c) *The set of Kronecker canonical forms of all pairs in a sufficiently small neighborhood of* (58) *consists of* (58) *and the pairs*

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

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

Let us denote by $\mathcal{P}^{\alpha_1\dots\alpha_m}_{\beta_1\dots\beta_n}$ the pair (59) in which $\alpha_1,\dots,\alpha_m,\beta_1,\dots,\beta_n$ are complex numbers.

Lemma 5.2. *If* $(C,D)=\mathcal{P}^{\alpha_1\dots\alpha_m}_{\beta_1\dots\beta_n}$*, then* $(D^T,C^T)\sim\mathcal{P}^{\beta_n\dots\beta_1}_{\alpha_m\dots\alpha_1}$*.*

Proof. We have the equivalences of pairs

$$\begin{aligned}
(D^T,C^T)&=\left(\left[\begin{array}{c|c} R_m & 0\\ \hline \begin{matrix}\beta_1 & & \\ \vdots & 0 & \\ \beta_n & & \end{matrix} & R_n^T\end{array}\right],\left[\begin{array}{c|c} L_m & 0\\ \hline \begin{matrix} & 0 & \\ \alpha_1 & \dots & \alpha_m\end{matrix} & L_n^T\end{array}\right]\right)\\
&\sim\left(\left[\begin{array}{c|c} R_n^T & \begin{matrix}\beta_1 & \\ \vdots & 0\\ \beta_n & \end{matrix}\\ \hline 0 & R_m\end{array}\right],\left[\begin{array}{c|c} L_n^T & \begin{matrix} & 0 & \\ \alpha_1 & \dots & \alpha_m\end{matrix}\\ \hline 0 & L_m\end{array}\right]\right)\\
&\sim\left(\left[\begin{array}{c|c} L_n^T & \begin{matrix} & \beta_n\\ 0 & \vdots\\ & \beta_1\end{matrix}\\ \hline 0 & L_m\end{array}\right],\left[\begin{array}{c|c} R_n^T & \begin{matrix}\alpha_m & \dots & \alpha_1\\ & 0 & \end{matrix}\\ \hline 0 & R_m\end{array}\right]\right)=\mathcal{P}^{\beta_n\dots\beta_1}_{\alpha_m\dots\alpha_1},
\end{aligned}$$

in which 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. □

Proof of Theorem 5.6. (a)&(b) By Theorem 3.2, there is a neighborhood of (58), in which each pair is equivalent to the pair

$$\mathcal{P}^{\alpha_1\dots\alpha_m}_{\beta_1\dots\beta_n} \tag{62}$$

for some $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n$.

Case 1: $\alpha_m \neq 0$ in (62). In this case, $\mathcal{P}^{\alpha_1 \ldots \alpha_m}_{\beta_1 \ldots \beta_n} \sim (I_{m+n-1}, \Phi)$ with

$$\Phi := \begin{bmatrix} -c_{m+n-2} & \cdots & -c_1 & -c_0 \\ 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \end{bmatrix}, \tag{63}$$

$$(c_0 \ldots, c_{m+n-2}) := \alpha_m^{-1}(-\beta_1, \ldots, -\beta_n, \alpha_1, \ldots, \alpha_{m-1}) \tag{64}$$

because

$$\mathcal{P}^{\alpha_1 \ldots \alpha_m}_{\beta_1 \ldots \beta_n}(Q_{m-1} \oplus Z_n) = (Q_m \oplus Z_{n-1})(I_{m+n-1}, \Phi), \tag{65}$$

in which

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

For example, if $m = n = 4$, then (65) 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 Jordan canonical form of (63) is $J_{r_1}(\lambda_1) \oplus \cdots \oplus J_{r_s}(\lambda_s)$ with distinct $\lambda_1, \ldots, \lambda_s \in \mathbb{C}$; its characteristic polynomial is

$$\begin{aligned} &(x-\lambda_1)^{r_1} \cdots (x-\lambda_s)^{r_s} = c_0 + c_1 x + \cdots + c_{m+n-2} x^{m+n-2} + x^{m+n-1} \\ &= \alpha_m^{-1}(-\beta_1 - \beta_2 x - \cdots - \beta_n x^{n-1} + \alpha_1 x^n + \alpha_2 x^{n+1} + \cdots + \alpha_m x^{m+n-1}). \end{aligned} \tag{66}$$

We have proved that

$$\mathcal{P}^{\alpha_1 \ldots \alpha_m}_{\beta_1 \ldots \beta_n} \sim (I, \Phi) \sim \mathcal{D}_{r_1}(\lambda_1) \oplus \cdots \oplus \mathcal{D}_{r_s}(\lambda_s) \qquad \text{if } \alpha_m \neq 0, \tag{67}$$

which is a pair of the form (61) with $l = 0$. This proves the statement (a) in Case 1.

By (67), each pair (61) with distinct eigenvalues $\lambda_1, \ldots, \lambda_s \in \mathbb{C}$ and $l = 0$ is equivalent to $\mathcal{P}^{\alpha_1 \ldots \alpha_m}_{\beta_1 \ldots \beta_n}$ defined by (66). Since (64) holds, $\mathcal{P}^{\alpha_1 \ldots \alpha_m}_{\beta_1 \ldots \beta_n}$ is the pair (59) defined by (60) with $\varepsilon = \alpha_m$. The pair $\mathcal{P}^{\alpha_1 \ldots \alpha_m}_{\beta_1 \ldots \beta_n}$ is also equivalent to the pair (59) defined by (60) with an arbitrary nonzero ε since

$$\left(\begin{bmatrix} L_m^T & P \\ 0 & L_n \end{bmatrix}, \begin{bmatrix} R_m^T & Q \\ 0 & R_n \end{bmatrix}\right) \begin{bmatrix} I_{m-1} & 0 \\ 0 & \delta I_n \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ 0 & \delta I_{n-1} \end{bmatrix} \left(\begin{bmatrix} L_m^T & \delta P \\ 0 & L_n \end{bmatrix}, \begin{bmatrix} R_m^T & \delta Q \\ 0 & R_n \end{bmatrix}\right) \tag{68}$$

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 (62). Let us show that

$$\mathcal{P}^{\alpha_1 \ldots \alpha_m}_{\beta_1 \ldots \beta_n} = \mathcal{P}^{\alpha_1 \ldots \alpha_k\, 0 \ldots 0}_{\beta_1 \;\ldots\; \beta_n} \sim \mathcal{P}^{\alpha_1 \ldots \alpha_k}_{\beta_1 \ldots \beta_n} \oplus (J_{m-k}(0), I_{m-k}) \qquad \text{if } \alpha_k \neq 0. \tag{69}$$

For clarity, we first prove (69) in the following special case:

$$\mathcal{P}^{\alpha_1\,\alpha_2\,0\,0}_{\beta_1\beta_2\beta_3\beta_4} \sim \mathcal{P}^{\alpha_1\,\alpha_2}_{\beta_1\beta_2\beta_3\beta_4} \oplus (J_2(0), I_2) \qquad \text{if } \alpha_2 \neq 0. \tag{70}$$

The first pair in (70) 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

Please cite this article in press as: V. Futorny et al., Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

$$(C,D)\sim\left(\left[\begin{array}{ccc|cccc} 1&0&0&&&&\alpha_1\\ 0&0&0&&&&\alpha_2\\ 0&0&1&&&&0\\ 0&0&0&&&&0\\ \hline &&&1&0&0&0\\ &&&0&1&0&0\\ &&&0&0&1&0 \end{array}\right],\left[\begin{array}{ccc|cccc} 0&0&0&\beta_1&\beta_2&\beta_3&\beta_4\\ 1&0&0&&&&\\ 0&1&0&&&&\\ 0&0&1&&&&\\ \hline &&&0&1&0&0\\ &&&0&0&1&0\\ &&&0&0&0&1 \end{array}\right]\right) \tag{71}$$

since 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 follows.
- 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 (71), which proves (70).

The equivalence (69) 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 $1, \dots, k-1$. 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 (69).

By (69) and Case 1,

$$\mathcal{P}^{\alpha_1 \dots \alpha_m}_{\beta_1 \dots \beta_n} = \mathcal{P}^{\alpha_1 \dots \alpha_k \, 0 \dots 0}_{\beta_1 \;\dots\; \beta_n} \sim \mathcal{D}_{r_1}(\lambda_1) \oplus \dots \oplus \mathcal{D}_{r_s}(\lambda_s) \oplus \mathcal{D}_{m-k}(\infty),$$

$$\alpha_k^{-1}(-\beta_1 - \dots - \beta_n x^{n-1} + \alpha_1 x^n + \dots + \alpha_k x^{n+k-1}) = \prod_{i=1}^{s}(x-\lambda_i)^{r_i}.$$

This proves the statement (a) in Case 2. This also proves the statement (b) in Case 2 for $\varepsilon = \alpha_k$; it holds for an arbitrary nonzero ε due to (68).

Case 3: $\alpha_1 = \dots = \alpha_m = 0$ in (62); that is, $(C, D) = \mathcal{P}^{0 \;\dots\; 0}_{\beta_1 \dots \beta_n}$. Let

$$\beta_1 = \dots = \beta_{p-1} = 0 \neq \beta_p, \quad \beta_q \neq 0 = \beta_{q+1} = \dots = \beta_n \quad (1 \leqslant p \leqslant q \leqslant n).$$

By Lemma 5.2 and Case 2, we have

$$\begin{aligned}(D^T, C^T) &\sim \mathcal{P}^{\beta_n \dots \beta_1}_{0 \;\dots\; 0} = \mathcal{P}^{0 \dots 0 \, \beta_q \dots \beta_p \, 0 \dots 0}_{0 \;\dots\; 0} \\ &\sim \mathcal{D}_{r_1}(\mu_1) \oplus \dots \oplus \mathcal{D}_{r_{s-1}}(\mu_{s-1}) \oplus \mathcal{D}_{m+n-q}(0) \oplus \mathcal{D}_{p-1}(\infty),\end{aligned}$$

in which $\mu_1, \dots, \mu_{s-1}$ are distinct nonzero complex numbers; this direct sum is determined by

$$(x-\mu_1)^{r_1} \cdots (x-\mu_{s-1})^{r_{s-1}} x^{m+n-q} = \beta_p^{-1}(\beta_q + \beta_{q-1}x + \dots + \beta_p x^{q-p}) x^{m+n-q}. \tag{72}$$

If $p < q$, we set $\lambda_1 := \mu_1^{-1}, \; \dots, \; \lambda_{s-1} := \mu_{s-1}^{-1}$, and find that

$$(C, D) \sim \mathcal{D}_{r_1}(\lambda_1) \oplus \dots \oplus \mathcal{D}_{r_{s-1}}(\lambda_{s-1}) \oplus \mathcal{D}_{m+n-q}(\infty) \oplus \mathcal{D}_{p-1}(0). \tag{73}$$

Replacing x by x^{-1} in the polynomials (72), we obtain consistently

$$\begin{aligned}(x^{-1} - \lambda_1^{-1})^{r_1} \cdots (x^{-1} - \lambda_{s-1}^{-1})^{r_{s-1}} &= \beta_p^{-1}(\beta_q + \beta_{q-1}x^{-1} + \dots + \beta_p x^{-(q-p)}), \\ (x-\lambda_1)^{r_1} \cdots (x-\lambda_{s-1})^{r_{s-1}} &= \beta_q^{-1}(\beta_p + \beta_{p+1}x + \dots + \beta_q x^{q-p}), \\ (x-\lambda_1)^{r_1} \cdots (x-\lambda_{s-1})^{r_{s-1}} x^{p-1} &= -\beta_q^{-1}(-\beta_p x^{p-1} - \beta_{p+1}x^p - \dots - \beta_q x^{q-1}).\end{aligned}$$

The last equality is the second equality in (60) that is determined by the first equality in (60) with $\varepsilon = -\beta_q$. The direct sum (73) is the direct sum (61). This proves the statement (a) in Case 3. This also proves the statement (b) in Case 3 for $\varepsilon = -\beta_q$; it holds for an arbitrary nonzero ε due to (68).

(c) This statement follows from (a), (b), Theorem 3.2, and Remark 3.1. □

Please cite this article in press as: V. Futorny et al., Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

6. Perturbations of Jordan matrices

By Lipschitz's property (see [38] or [1]), 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 with a nonsingular first matrix, with respect to equivalence transformations, is reduced to the theory of perturbations of square matrices with respect to similarity. By Theorem 4.1, it reduces to the theory of perturbations of Jordan matrices with a single eigenvalue.

The closures of orbits of Jordan matrices under similarity have been described by Den Boer and Thijsse [13], and by Markus and Parilis [35]; see also [22, Theorem 2.1]. In this section, we describe the closures of orbits of Jordan matrices in the form that is used in Theorem I. The proof is based on the Weyr canonical form of matrices under similarity.

Theorem 6.1. *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 two Jordan blocks. Write it as follows:*

$$J = P \oplus J_p(\lambda) \oplus J_q(\lambda) \oplus Q \quad \text{for some } p \leqslant q, \tag{74}$$

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 Q can be absent). Define the Jordan matrix

$$J_{p,q} := P \oplus J_{p-1}(\lambda) \oplus J_{q+1}(\lambda) \oplus Q, \tag{75}$$

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.

Lemma 6.1. *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 (see Section 3.3). Write*

$$s_i := m_1 + \dots + m_i \quad \text{and} \quad s'_i := m'_1 + \dots + m'_i \quad \text{for all } i. \tag{76}$$

Then

$$\langle J\rangle \leqslant \langle J'\rangle \quad \Longleftrightarrow \quad s_i \geqslant s'_i \text{ for all } i. \tag{77}$$

Please cite this article in press as: V. Futorny et al., Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

Example 6.1. 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, $\langle J \rangle \leqslant \langle J' \rangle$, $\langle J \rangle \ngeqslant \langle J' \rangle$, and so $\langle J \rangle < \langle J' \rangle$.

Proof of Lemma 6.1. Let J be a Jordan matrix with a single eigenvalue λ. Then $\langle J - \lambda I \rangle = \langle J \rangle - \lambda I$ and we must prove (77) only for $\lambda = 0$.

$\Longleftarrow$. Let W and W' be Weyr canonical matrices of the same size with the single eigenvalue 0 (see Section 3.3). Let their Weyr characteristics $(m_1, m_2, \dots, m_k)$ and $(m'_1, m'_2, \dots, m'_k)$ satisfy $s_1 \geqslant s'_1$, $s_2 \geqslant s'_2, \dots$. Then 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 $\langle W \rangle \leqslant \langle W' \rangle$.

$\Longrightarrow$. Let J be a Jordan matrix with the single eigenvalue $\lambda = 0$. 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

$$A_1, A_2, \dots \to J \tag{78}$$

in which all A_i are similar to J'. All A_i have the same characteristic polynomial $f(x)$. Since the coefficients of 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 obtain a unitary matrix $U_i = S_i R_i$, where R_i is a nonsingular upper-triangular matrix. Then

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

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 (78) if necessarily, we obtain $U_i \to U$. Since each U_i is unitary, we have $U_iU_i^* = I$, and so $UU^* = I$. Hence U is unitary and

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

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

Therefore,

$$m_1 = \operatorname{nullity} J = \operatorname{nullity} U^{-1}JU \geqslant m_1'.$$

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

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

and so on, which proves "$\Longrightarrow$" in (77). □

Proof of Theorem 6.1. (a) Let $J = J_p(\lambda)$ and $\langle J\rangle \leqslant \langle J'\rangle$. By (77), $m_1' \leqslant m_1 = 1$. Since m_1' is the number of Jordan blocks, J' is a Jordan block. Since J and J' have the same size, $J' = J_p(\lambda) = J$.

(b) For each matrix X, we denote by $(m_1(X), m_2(X), \dots)$ its Weyr characteristic and write $s_i(X) := m_1(X) + \dots + m_i(X)$. 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 (77),

$$\langle A \oplus B\rangle \leqslant \langle A \oplus C\rangle \quad \Longleftrightarrow \quad \langle B\rangle \leqslant \langle C\rangle. \tag{79}$$

Let $(m_1, m_2, \dots)$ and $(\widetilde{m}_1, \widetilde{m}_2, \dots)$ be the Weyr characteristics of the matrices (74) and (75). Then $\widetilde{m}_p = m_p - 1$, $\widetilde{m}_{q+1} = m_{q+1} + 1$, the other $\widetilde{m}_i = m_i$, and so

$$\widetilde{s}_p = s_p - 1,\ \widetilde{s}_{p+1} = s_{p+1} - 1,\ \dots,\ \widetilde{s}_q = s_q - 1, \quad \text{the other } \widetilde{s}_i = s_i \tag{80}$$

in the notation (76). Let us prove three facts.

Fact 1: $\langle J\rangle < \langle J_{p,q}\rangle$. This inequality follows from (79) and the inequality $\langle J_p(\lambda) \oplus J_q(\lambda)\rangle < \langle J_{p-1}(\lambda) \oplus J_{q+1}(\lambda)\rangle$, which holds by (77) and (80).

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

$$\langle J\rangle < \langle J'\rangle \quad \Longrightarrow \quad \langle J\rangle < \langle J_{p,q}\rangle \leqslant \langle J'\rangle \textit{ for some } p, q. \tag{81}$$

Due to (79), it is sufficient to prove (81) for J and J' that have no common Jordan blocks. By the assumptions of Theorem 6.1(b), J has at least two Jordan blocks. Let us show that (81) holds for p and q 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 $\geqslant q$.

By $\langle J'\rangle \geqslant \langle J\rangle$ and Lemma 6.1, $s'_i \leqslant s_i$ for all i. By Fact 1, $\langle J_{p,q}\rangle > \langle J\rangle$. We must prove that $\langle J'\rangle \geqslant \langle J_{p,q}\rangle$; i.e., $s'_i \leqslant \widetilde{s}_i$ for all i. Due to (80), it suffices to prove that

$$s'_p < s_p, \quad s'_{p+1} < s_{p+1}, \quad \dots, \quad s'_q < s_q. \tag{82}$$

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

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

Thus, $m_p \geqslant m'_p$.

If $m_p = m'_p$, then

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

Hence, $s_1 = s'_1$, $s_2 = s'_2$, …, $s_p = s'_p$, $s_{p+1} = s_p + m_{p+1} < s'_p + m'_{p+1} = s'_{p+1}$, 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 (82) 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

$$\begin{array}{l} m_p - 1 = m_{p+1} = m_{p+2} = \cdots = m_q \\ \quad\;\;\, \mathbin{\rotatebox[origin=c]{90}{$\geqslant$}} \\ \quad m'_p \;\; = m'_{p+1} \geqslant m'_{p+2} \geqslant \cdots \geqslant m'_q \end{array}$$

We obtain consistently $s_p > s'_p$, $s_{p+1} = s_p + m_{p+1} > s'_p + m'_{p+1} = s'_{p+1}$, …, $s_q = s_{q-1} + m_q > s'_{q-1} + m'_q = s'_q$, which proves (82) 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 $n(J) := t_1 + t_2 + \cdots$ is the number of Jordan blocks in J. If $t = (t_1, t_2, \dots)$ and $t' = (t'_1, t'_2, \dots)$ are infinite sequences of nonnegative integers with $t_1 + t_2 + \cdots < \infty$ and $t'_1 + t'_2 + \cdots < \infty$, then we write

$$\begin{aligned} t' \overset{l}{\prec} t \quad & \text{if } t'_1 = t_1,\ \dots,\ t'_{k-1} = t_{k-1},\ t_k < t_k \text{ for some } k \geqslant 1; \\ t' \overset{r}{\prec} t \quad & \text{if } t'_k < t_k,\ t'_{k+1} = t_{k+1},\ t'_{k+2} = t_{k+2}, \dots \text{ for some } k \geqslant 1. \end{aligned}$$

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 $J \downarrow J_{s,r}$:

$$J_{p',q'} \downarrow (J_{p',q'})_{r_1,s_1} \downarrow ((J_{p',q'})_{r_1,s_1})_{r_2,s_2} \downarrow \cdots \downarrow J_{p,q}. \tag{83}$$

Therefore,

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

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

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

since the analogous statements hold for each of the replacements (83).

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

Thus, $n(J_{p',q'}) = n(J_{p,q})$. 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 $(p', q') \neq (p, q)$ and proves Fact 3.

Facts 2 and 3 prove Theorem 6.1(b). □

The following theorem ensures Theorem 5.5.

Theorem 6.2.

(a) *Let $\alpha_1, \dots, \alpha_m$ be complex numbers that are not all zero. Let α_r be the last nonzero element in this sequence. Then the matrix*

$$\left[\begin{array}{c|cc} & \alpha_1 & \\ J_m(\lambda) & \vdots & 0 \\ & \alpha_m & \\ \hline 0 & & J_n(\lambda) \end{array}\right], \qquad m \leqslant n, \;\; \lambda \in \mathbb{C} \tag{84}$$

is similar to

$$J_{m-r}(\lambda) \oplus J_{n+r}(\lambda). \tag{85}$$

(b) *If a given Jordan matrix J is similar to a matrix in an arbitrarily small neighborhood of*

$$J_m(\lambda) \oplus J_n(\lambda), \qquad m \leqslant n, \;\; \lambda \in \mathbb{C},$$

then J has the form $J_{m-r}(\lambda) \oplus J_{n+r}(\lambda)$ with $r \in \{0, \dots, m\}$.

Proof. (a) Let A be the matrix (84). Using similarity transformations

$$A \mapsto \begin{bmatrix} K_m & 0 \\ 0 & I_n \end{bmatrix} A \begin{bmatrix} K_m^{-1} & 0 \\ 0 & I_n \end{bmatrix}, \quad K_m := \begin{bmatrix} \kappa_1 & \kappa_2 & \ddots & \kappa_m \\ & \kappa_1 & \ddots & \ddots \\ & & \ddots & \kappa_2 \\ 0 & & & \kappa_1 \end{bmatrix}, \quad \kappa_1 \neq 0,$$

we make $(\alpha_1, \dots, \alpha_n) = (0, \dots, 0, 1, 0, \dots, 0)$ with "1" at the position r, preserving the other entries of A (compare with (49) and (50)). In the obtained matrix

Please cite this article in press as: V. Futorny et al., Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

$$
\begin{array}{cc|cc|l}
 & \scriptstyle r+1 \qquad\quad m & \scriptstyle 1 \qquad\quad m-r & & \\
\hline
\lambda\ \ 1 & & \vdots & & \scriptstyle 1\\
\ \ \ddots\ \ \ddots & & 0 & & \\
\qquad\quad \lambda & 1 & 1 & & \scriptstyle r\\
 & \lambda\ \ 1 & 0 & & \scriptstyle r+1\\
 & \quad\lambda\ \ \ddots & 0 & & \\
 & \qquad\ddots\ \ 1 & \vdots & & \\
 & \qquad\quad \lambda & 0 & & \scriptstyle m\\
\hline
 & 0\ \ \not{0} & \lambda\ \ 1 & & \scriptstyle 1\\
 & \quad 0\ \ \ddots & \quad\lambda\ \ \ddots & & \\
 & \qquad\ddots\ \ \not{0} & \qquad\ddots\ \ 1 & & \\
 & \qquad\quad 0 & \qquad\quad \lambda & 1 & \scriptstyle m-r\\
 & & & \lambda\ \ \ddots & \\
 & & & \quad\ddots\ \ 1 & \\
 & & & \qquad\quad\lambda & \scriptstyle n
\end{array}
\tag{86}
$$

we make zero the entry “1” to the left of $\alpha_r = 1$ by the following similarity transformations (every $\not{0}$ denotes the zero entry that first is transformed to -1 and then is restored to 0; compare with (37)):

- Make zero the entry “1” to the left of $\alpha_r = 1$ by subtracting columns $1, 2, \dots, m-r$ of the second vertical strip from 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 (86) is subtracted from the marked $(m-r) \times (m-r)$ subblock in the $(2,1)$th block.
- Make the inverse transformations of rows, adding rows $r+1, \dots, m$ of the first horizontal strip to rows $1, \dots, m-r$ of the second horizontal strip, restoring the $(m-r) \times (m-r)$ subblock in the $(2,1)$th block.

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 (85).

(b) This statement follows from Theorem 6.1(b). □

7. Theorem II follows from Theorem I

Theorem II is formulated in terms of coin moves and proved sketchily by Edelman, Elmroth, and Kågström [22, Theorem 3.2]. In this section we derive Theorem II from Theorem I.

It is sufficient to prove the following statement:

$$\text{\textit{Let a Kronecker pair } } \mathcal{B} \text{\textit{ be obtained from a Kronecker pair }} \mathcal{A} \text{\textit{ by some replacement }} (\mathrm{j}) \text{\textit{ from Theorem I, where }} \mathrm{j} \in \{\mathrm{i}, \mathrm{ii}, \dots, \mathrm{vi}\}. \text{\textit{ Then }} \langle\mathcal{B}\rangle \text{\textit{ immediately succeeds }} \langle\mathcal{A}\rangle \text{\textit{ if and only if }} (\mathrm{j}) \text{\textit{ is the replacement }} (\mathrm{j}') \text{\textit{ from Theorem II.}} \tag{87}$$

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

$$\mathcal{L}_m \oplus \mathcal{L}_n \downarrow \mathcal{L}_{m+1} \oplus \mathcal{L}_{n-1}, \quad \text{in which } m+2 \leqslant n. \tag{88}$$

$\Longrightarrow$. Let $\langle\mathcal{B}\rangle$ immediately succeed $\langle\mathcal{A}\rangle$. We must prove that (88) is the replacement (ii′). To the contrary, let (88) be not the replacement (ii′); that is, $\mathcal{A}$ contains $\mathcal{L}_m \oplus \mathcal{L}_k \oplus \mathcal{L}_n$ with $m < k < n$ and $n - m \geqslant 3$.

If $k - m \geqslant 2$, then (88) is the following composition of replacements of type (ii):

$$\mathcal{L}_m \oplus \mathcal{L}_k \oplus \mathcal{L}_n \downarrow \mathcal{L}_{m+1} \oplus \mathcal{L}_{k-1} \oplus \mathcal{L}_n \downarrow \mathcal{L}_{m+1} \oplus \mathcal{L}_k \oplus \mathcal{L}_{n-1}.$$

By Theorem I,

$$\langle \mathcal{L}_m \oplus \mathcal{L}_k \oplus \mathcal{L}_n \rangle < \langle \mathcal{L}_{m+1} \oplus \mathcal{L}_{k-1} \oplus \mathcal{L}_n \rangle < \langle \mathcal{L}_{m+1} \oplus \mathcal{L}_k \oplus \mathcal{L}_{n-1} \rangle,$$

and so $\langle\mathcal{B}\rangle$ is not an immediate successor of $\langle\mathcal{A}\rangle$.

If $k - m = 1$, then $n - k \geqslant 2$ and (88) is the following composition of replacements of type (ii):

$$\mathcal{L}_m \oplus \mathcal{L}_k \oplus \mathcal{L}_n \downarrow \mathcal{L}_m \oplus \mathcal{L}_{k+1} \oplus \mathcal{L}_{n-1} \downarrow \mathcal{L}_{m+1} \oplus \mathcal{L}_k \oplus \mathcal{L}_{n-1}.$$

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 (ii′). Let $\mathcal{B}$ can be also obtained from $\mathcal{A}$ by a sequence

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

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 $p = 1$.

Let

$$\begin{gathered}\mathcal{A} = \bigoplus_{i=1}^{\underline{s}} \mathcal{L}_{m_i}^T \oplus \bigoplus_{i=1}^{\overline{s}} \mathcal{L}_{n_i} \oplus \bigoplus_{i=1}^{t} \Big(\mathcal{D}_{k_{i1}}(\lambda_i) \oplus \cdots \oplus \mathcal{D}_{k_{is_i}}(\lambda_i)\Big), \\ m_1 \leqslant \cdots \leqslant m_{\underline{s}}, \quad n_1 \leqslant \cdots \leqslant n_{\overline{s}}, \quad k_{i1} \leqslant \cdots \leqslant k_{is_i} \ (i = 1, \ldots, t),\end{gathered} \tag{89}$$

in which $\lambda_1, \ldots, \lambda_t \in \mathbb{C} \cup \infty$ are distinct (see (16)).

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

- type (vi) since $\mathcal{A}$ and $\mathcal{B}$ have the same number $\underline{s}$ of summands of type $\mathcal{L}_m^T$, but (vi) decreases the number $\underline{s}$ and this number cannot be restored by (i)–(v);
- type (iii) since it increases the number $m_1 + \cdots + m_{\underline{s}}$ whereas this number is not changed by (i), (ii), (iv), and (v);
- type (iv) since it increases $n_1 + \cdots + n_{\overline{s}}$;

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

Therefore, all $\varphi_1, \dots, \varphi_p$ are replacements of type (ii). Since each replacement (ii′) is not the composition of several replacements of type (ii), $p = 1$, and so $\langle \mathcal{B} \rangle$ is an immediate successor of $\langle \mathcal{A} \rangle$. We have proved (87) in Case 1.

Case 2: (j) *is the replacement* (i). The statement (87) is proved by transposing the matrices in Case 1.

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

$$\mathcal{L}_m \oplus \mathcal{D}_n(\lambda) \downarrow \mathcal{L}_{m+1} \oplus \mathcal{D}_{n-1}(\lambda). \tag{90}$$

$\Longrightarrow$. To the contrary, suppose that (90) is not (iv′); that is, $m < \overline{m}$ or $n < \overline{n}_\lambda$. If $m < \overline{m}$, then (90) is the composition of replacements of types (ii) and (iv):

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

If $n < \overline{n}_\lambda$, then

$$\begin{aligned} \mathcal{L}_m \oplus \mathcal{D}_n(\lambda) \oplus \mathcal{D}_{\overline{n}_\lambda}(\lambda) &\downarrow \mathcal{L}_{m+1} \oplus \mathcal{D}_n(\lambda) \oplus \mathcal{D}_{\overline{n}_\lambda - 1}(\lambda) \\ &\downarrow \mathcal{L}_{m+1} \oplus \mathcal{D}_{n-1}(\lambda) \oplus \mathcal{D}_{\overline{n}_\lambda}(\lambda). \end{aligned}$$

By Theorem I, $\langle \mathcal{B} \rangle$ is not an immediate successor of $\langle \mathcal{A} \rangle$.

$\Longleftarrow$. Let $\mathcal{B}$ be obtained from $\mathcal{A}$ by replacement (iv′). Let $\mathcal{B}$ can 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), which decreases the number $\underline{s}$ (see (89));
- type (ii), which increases lexicographically $(n_1, n_2, \dots, n_{\underline{s}})$;
- types (i) and (iii), which change the sequence $(m_1, m_2, \dots, m_{\overline{s}})$;
- type (v), which decreases lexicographically $(k_{i1}, k_{i2}, \dots, k_{is_i})$ if $\lambda_i = \lambda$.

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

Case 4: (j) *is the replacement* (iii). The statement (87) is proved by transposing the matrices in Case 3.

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

$$\mathcal{D}_m(\lambda) \oplus \mathcal{D}_n(\lambda) \downarrow \mathcal{D}_{m-1}(\lambda) \oplus \mathcal{D}_{n+1}(\lambda), \text{ in which } m \leqslant n. \tag{91}$$

Please cite this article in press as: V. Futorny et al., Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

$\Longrightarrow$. To the contrary, suppose that (91) is not (v′); that is, $\mathcal{A}$ contains $\mathcal{D}_m(\lambda)\oplus\mathcal{D}_k(\lambda)\oplus\mathcal{D}_n(\lambda)$ with $m \leqslant k \leqslant n$ and $m < n$. Then $\langle\mathcal{B}\rangle$ is not an immediate successor of $\langle\mathcal{A}\rangle$ since if $m \leqslant k < n$ then

$$\begin{aligned}\mathcal{D}_m(\lambda)\oplus\mathcal{D}_k(\lambda)\oplus\mathcal{D}_n(\lambda) &\rightsquigarrow \mathcal{D}_{m-1}(\lambda)\oplus\mathcal{D}_{k+1}(\lambda)\oplus\mathcal{D}_n(\lambda)\\ &\rightsquigarrow \mathcal{D}_{m-1}(\lambda)\oplus\mathcal{D}_k(\lambda)\oplus\mathcal{D}_{n+1}(\lambda),\end{aligned}$$

and if $m < k \leqslant n$ then

$$\begin{aligned}\mathcal{D}_m(\lambda)\oplus\mathcal{D}_k(\lambda)\oplus\mathcal{D}_n(\lambda) &\rightsquigarrow \mathcal{D}_m(\lambda)\oplus\mathcal{D}_{k-1}(\lambda)\oplus\mathcal{D}_{n+1}(\lambda)\\ &\rightsquigarrow \mathcal{D}_{m-1}(\lambda)\oplus\mathcal{D}_k(\lambda)\oplus\mathcal{D}_{n+1}(\lambda).\end{aligned}$$

$\Longleftarrow$. Let $\mathcal{B}$ be obtained from $\mathcal{A}$ by replacement (v′), and let $\mathcal{B}$ can 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 types (i)–(iv) and (vi) since they change $m_1,\dots,m_{\underline{s}}$ or $n_1,\dots,n_{\overline{s}}$ (see (89)).

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)*:*

$$\mathcal{L}_m^T\oplus\mathcal{L}_n \rightsquigarrow \mathcal{D}_{r_1}(\mu_1)\oplus\cdots\oplus\mathcal{D}_{r_q}(\mu_q), \tag{92}$$

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

$\Longrightarrow$. To the contrary, suppose that (92) is not (vi′).

If $m<\overline{m}$, then

$$\begin{aligned}\mathcal{L}_m^T\oplus\mathcal{L}_{\overline{m}}^T\oplus\mathcal{L}_n &\rightsquigarrow \mathcal{L}_m^T\oplus\mathcal{D}_{r_1+\overline{m}-m}(\mu_1)\oplus\cdots\oplus\mathcal{D}_{r_q}(\mu_q)\\ &\rightsquigarrow \mathcal{L}_{\overline{m}}^T\oplus\mathcal{D}_{r_1}(\mu_1)\oplus\cdots\oplus\mathcal{D}_{r_q}(\mu_q),\end{aligned}$$

and so $\langle\mathcal{B}\rangle$ is not an immediate successor of $\langle\mathcal{A}\rangle$. Hence $m=\overline{m}$ and, analogously, $n=\overline{n}$.

If some $\lambda_i\notin\{\mu_1,\dots,\mu_q\}$ (see (89)), then

$$\begin{aligned}\mathcal{L}_{\overline{m}}^T\oplus\mathcal{L}_{\overline{n}}\oplus\mathcal{D}_{k_{i1}}(\lambda_i) &\rightsquigarrow \mathcal{L}_{\overline{m}}^T\oplus\mathcal{L}_{\overline{n}+k_{i1}}\\ &\rightsquigarrow \mathcal{D}_{r_1}(\mu_1)\oplus\cdots\oplus\mathcal{D}_{r_q}(\mu_q)\oplus\mathcal{D}_{k_{i1}}(\lambda_i),\end{aligned}$$

and so $\langle\mathcal{B}\rangle$ is not an immediate successor of $\langle\mathcal{A}\rangle$. Hence $q\geqslant t$ (see (89)) 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 $\mu_1=\lambda_1$,

$$\begin{aligned}\mathcal{L}_{\overline{m}}^T\oplus\mathcal{L}_{\overline{n}}\oplus\mathcal{D}_{k_{1s_1}}(\mu_1) &\rightsquigarrow \mathcal{L}_{\overline{m}}^T\oplus\mathcal{L}_{\overline{n}+k_{1s_1}-r_1}\oplus\mathcal{D}_{r_1}(\mu_1)\\ &\rightsquigarrow \mathcal{D}_{r_2}(\mu_2)\oplus\cdots\oplus\mathcal{D}_{r_q}(\mu_q)\oplus\mathcal{D}_{k_{1s_1}}(\mu_1)\oplus\mathcal{D}_{r_1}(\mu_1),\end{aligned}$$

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

$$\varphi: \ \mathcal{L}^T_{m_{\underline{s}}} \oplus \mathcal{L}_{n_{\overline{s}}} \downarrow \mathcal{D}_{r_1}(\mu_1) \oplus \cdots \oplus \mathcal{D}_{r_q}(\mu_q), \qquad q \geqslant t \tag{93}$$

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 $\mathcal{B}$ can 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 increases $\sum k_{ij}$ and decreases $\underline{s}$ by one. The preceding replacements $\varphi_1, \ldots, \varphi_{u-1}$ do not change $\underline{s}$ and $\overline{s}$. 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 (vi′) due to part "$\Longrightarrow$"; that is,

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

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 (iii). However, this φ_l decreases $\sum k_{ij}$, which cannot be restored because of the condition $k_{1s_1} \leqslant r_1, \ldots, k_{ts_t} \leqslant r_t$. Hence $m'_{\underline{s}} \leqslant m_{\underline{s}}$. 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 (93). It is easy to check that $u = p = 1$ and $\varphi_1 = \varphi$.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgements

The authors wish to express their gratitude to the referee for several helpful comments. V. Futorny was supported by the CNPq (304467/2017-0) and the FAPESP (2018/23690-6). V.V. Sergeichuk is greatly indebted to the University of Sao Paulo, where the paper was written, for hospitality and the FAPESP for financial support (2018/24089-4).

References

[1] A. Alazemi, M. Anđelić, 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] V.I. Arnold, On matrices depending on parameters, Russ. Math. Surv. 26 (2) (1971) 29–43.
[3] V.I. Arnold, Lectures on bifurcations in versal families, Russ. Math. Surv. 27 (5) (1972) 54–123.
[4] V.I. Arnold, Geometrical Methods in the Theory of Ordinary Differential Equations, Springer-Verlag, 1988.
[5] J. Bender, K. Bongartz, Minimal singularities in orbit closures of matrix pencils, Linear Algebra Appl. 365 (2003) 13–24.
[6] D.L. Boley, The algebraic structure of pencils and block Toeplitz matrices, Linear Algebra Appl. 279 (1998) 255–279.
[7] K. Bongartz, Minimal singularities for representations of Dynkin quivers, Comment. Math. Helv. 69 (1994) 575–611.
[8] K. Bongartz, Degenerations for representations of tame quivers, Ann. Sci. Éc. Norm. Supér. (4) 28 (1995) 647–688.
[9] K. Bongartz, On degenerations and extensions of finite dimensional modules, Adv. Math. 121 (1996) 245–287.
[10] K. Bongartz, Some geometric aspects of representation theory, in: Algebras and Modules, I, Trondheim, 1996, in: CMS Conf. Proc., vol. 23, Amer. Math. Soc., Providence, RI, 1998, pp. 1–27.
[11] V.A. Bovdi, M.A. Salim, V.V. Sergeichuk, Neighborhood radius estimation for Arnold's miniversal deformations of complex and p-adic matrices, Linear Algebra Appl. 512 (2017) 97–112.
[12] F. De Terán, F.M. Dopico, A note on generic Kronecker orbits of matrix pencils with fixed rank, SIAM J. Matrix Anal. Appl. 30 (2008) 491–496.
[13] H. Den Boer, G.Ph.A. Thijsse, Semi-stability of sums of partial multiplicities under additive perturbation, Integral Equ. Oper. Theory 3 (1980) 23–42.
[14] A. Dmytryshyn, Miniversal deformations of pairs of skew-symmetric matrices under congruence, Linear Algebra Appl. 506 (2016) 506–534.
[15] A. Dmytryshyn, Miniversal deformations of pairs of symmetric matrices under congruence, Linear Algebra Appl. 568 (2019) 84–105.
[16] A. Dmytryshyn, F.M. Dopico, Generic skew-symmetric matrix polynomials with fixed rank and fixed odd grade, Linear Algebra Appl. 536 (2018) 1–18.
[17] A.R. Dmytryshyn, V. Futorny, V.V. Sergeichuk, Miniversal deformations of matrices of bilinear forms, Linear Algebra Appl. 436 (2012) 2670–2700.
[18] A. Dmytryshyn, V. Futorny, B. Kågström, L. Klimenko, V.V. Sergeichuk, Change of the congruence canonical form of 2-by-2 and 3-by-3 matrices under perturbations and bundles of matrices under congruence, Linear Algebra Appl. 469 (2015) 305–334.
[19] A.R. Dmytryshyn, V. Futorny, V.V. Sergeichuk, Miniversal deformations of matrices under *congruence and reducing transformations, Linear Algebra Appl. 446 (2014) 388–420.
[20] A. Dmytryshyn, S. Johansson, B. Kagstrom, P. Van Dooren, Geometry of matrix polynomial spaces, Found. Comput. Math. 20 (2020) 423–450.
[21] A. Edelman, E. Elmroth, B. Kågström, A geometric approach to perturbation theory of matrices and matrix pencils. Part I: versal deformations, SIAM J. Matrix Anal. Appl. 18 (1997) 653–692.
[22] 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.
[23] E. Elmroth, P. Johansson, B. Kågström, Computation and presentation of graphs displaying closure hierarchies of Jordan and Kronecker structures, Numer. Linear Algebra Appl. 8 (2001) 381–399.
[24] E. Elmroth, B. Kågström, The set of 2-by-3 matrix pencils — Kronecker structures and their transitions under perturbations, SIAM J. Matrix Anal. Appl. 17 (1996) 1–34.
[25] V. Futorny, L. Klimenko, V.V. Sergeichuk, Change of the *congruence canonical form of 2-by-2 matrices under perturbations, Electron. J. Linear Algebra 27 (2014) 146–154.
[26] 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 (2003) 943–962.
[27] 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.
[28] S. Johansson, Reviewing the Closure Hierarchy of Orbits and Bundles of System Pencils and Their Canonical Forms, Technical report, Umeå University, Department of Computing Science, 2009, UMINF-09.02.

Please cite this article in press as: V. Futorny et al., Perturbation theory of matrix pencils through miniversal deformations, Linear Algebra Appl. (2021), https://doi.org/10.1016/j.laa.2020.12.009

[29] B. Kågström, S. Johansson, P. Johansson, StratiGraph tool: matrix stratifications in control applications, in: L.T. Biegler, et al. (Eds.), Control and Optimization with Differential-Algebraic Constraints, in: Adv. Des. Control, vol. 23, SIAM, Philadelphia, PA, 2012, pp. 79–103.

[30] L. Klimenko, V.V. Sergeichuk, Block triangular miniversal deformations of matrices and matrix pencils, in: V. Olshevsky, E. Tyrtyshnikov (Eds.), Matrix Methods: Theory, Algorithms and Applications, World Sci. Publ., Hackensack, NJ, 2010, pp. 69–84.

[31] L. Klimenko, V.V. Sergeichuk, An informal introduction to perturbations of matrices determined up to similarity or congruence, São Paulo J. Math. Sci. 8 (2014) 1–22.

[32] A.A. Mailybaev, Reduction of matrix families to normal forms and its application to stability theory, Fundam. Prikl. Mat. 5 (1999) 1111–1133 (in Russian).

[33] 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.

[34] A.A. Mailybaev, Transformation to versal deformations of matrices, Linear Algebra Appl. 337 (2001) 87–108.

[35] A.S. Markus, E.È. Parilis, The change of the Jordan structure of a matrix under small perturbations, Mat. Issled. 54 (1980) 98–109 (in Russian), English translation: Linear Algebra Appl. 54 (1983) 139–152.

[36] K. O'Meara, J. Clark, C. Vinsonhaler, Advanced Topics in Linear Algebra: Weaving Matrix Problems Through the Weyr Form, Oxford University Press, 2011.

[37] A. Pokrzywa, On perturbations and the equivalence orbit of a matrix pencil, Linear Algebra Appl. 82 (1986) 99–121.

[38] L. Rodman, Remarks on Lipschitz properties of matrix groups actions, Linear Algebra Appl. 434 (2011) 1513–1524.

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

[40] StratiGraph and MCS Toolbox, Software Tools, Department of Computing Science, Umeå University, Sweden, [Website], www.cs.umu.se/english/research/groups/matrix-computations/stratigraph.