On Perturbations and the Equivalence Orbit of a Matrix Pencil

Andrzej Pokrzywa
Institute of Mathematics
Polish Academy of Sciences
Sniadeckich 8
00-950 Warsaw, P.O.B. 137, Poland

Submitted by Hans Schneider

ABSTRACT

Two theorems on the canonical Kronecker form of a perturbed matrix pencil and the characterization of the closure of the set of all matrix pencils with a fixed Kronecker canonical form are given.

INTRODUCTION

In the last few years a number of algorithms for computing the Kronecker canonical form of a singular matrix pencil have been developed (see the articles of V. N. Kublanovskaya and P. Van Dooren in [8]). The problem itself is in general ill posed, since the Kronecker canonical form may change under small perturbations of matrices defining a given matrix pencil. The aim of this paper is to study some aspects of this phenomenon.

Two theorems on the canonical Kronecker form of a perturbed pencil are stated. These theorems are obtained using simple tensor calculus. The application of this calculus to matrix pencils resembles in some ways the application of the functional calculus to operators. The perturbation theorems obtained here are related (in the author's opinion) to some perturbation results for Fredholm operators (cf. [5]).

It is shown also that theorems on perturbations of matrix pencils are in some sense the best possible. This aim is achieved by describing the closure of the equivalence orbit of a matrix pencil. This last concept is analogous to the concept of the similarity orbit of an operator, and there are some common aspects of Theorem 3 of this paper characterizing the closure of the equivalence orbit of a matrix pencil and a deep characterization of the closure of the similarity orbit of an operator given in [1, Theorem 1].

© Elsevier Science Publishing Co., Inc., 1986
52 Vanderbilt Ave., New York, NY 10017 0024-3795/86/$3.50
Let $X, Y$ be two complex linear finite dimensional spaces. A function $\mathcal{A}(\lambda) = A + \lambda B$, where $A, B$ are linear operators acting from $X$ to $Y$, is called an $(X, Y)$ operator pencil and is defined for all $\lambda \in \mathbb{C}$ if we put $A(\infty) = B$. Here $\mathbb{C}$ (the complex plane with the point at infinity) is a compact topological space, the $\varepsilon$-neighborhood of $\infty$ is defined as $\{ z \in \mathbb{C}; |z| > \varepsilon^{-1} \} \cup \{ \infty \}$, and the $\varepsilon$-neighborhoods of other points are defined in the usual way. An $(X, Y)$-pencil $\mathcal{A}$ is called decomposable if there exist two pairs $(X_1, Y_1), (X_2, Y_2)$ of subspaces such that $X = X_1 \oplus X_2$, $Y = Y_1 \oplus Y_2$, $(X_i, Y_i) \neq (X, Y)$, and $\mathcal{A}(\lambda)X_i \subset Y_i$ for all $\lambda \in \mathbb{C}$, $i = 1, 2$. A pencil which is not decomposable is called indecomposable. Each indecomposable pencil has one of the following matrix representations [1, Proposition 2.21, with appropriately chosen bases of $X$ and $Y$:]

$$
\mathcal{L}_j(\lambda) = \begin{pmatrix}
-\lambda & -\lambda & & \\
1 & -\lambda & & \\
& 1 & -\lambda & \\
& & 1 & -\lambda \\
\end{pmatrix}^{j+1}, \quad \mathcal{R}_j = (\mathcal{L}_j)^T,
$$

$$
\mathcal{A}_j(\mu, \lambda) = \begin{cases}
(\mu - \lambda)I_j + N_j, & \mu \in \mathbb{C} \\
I_j - \lambda N_j, & \mu = \infty
\end{cases}, \quad j = 1, 2, \ldots,
$$

where $I_j$ stands for the $j \times j$ identity matrix and

$$
N_j = \begin{pmatrix}
0 & 1 & & \\
0 & 1 & & \\
& 1 & & \\
& & \ddots & \\
& & & 1 \\
0 & & & \\
\end{pmatrix}^{j}.
$$

The Weierstrass-Kronecker theorem states that each $(X, Y)$-pencil $\mathcal{A}$ may be decomposed into a direct sum of indecomposable pencils, and although the
decomposition may be not unique, the numbers $l_j(\mathcal{A}), r_j(\mathcal{A}), d_j(\mu, \mathcal{A})$ of pencils of types $\mathcal{L}_j, \mathcal{R}_j, \mathcal{I}_j(\mu)$ (respectively) which appear in the decomposition of $\mathcal{A}$ do not depend on this decomposition. We shall refer to these numbers as the Kronecker structure of $\mathcal{A}$. We put also $l(\mathcal{A}) = \sum_{j \geq 0} l_j(\mathcal{A})$ and $r(\mathcal{A}) = \sum_{j \geq 0} r_j(\mathcal{A})$. Note that for each $(X, Y)$-pencil $\mathcal{A}$ the following identities hold:

$$\dim X = r(\mathcal{A}) + \sum_{j \geq 0} j \left( l_j(\mathcal{A}) + r_j(\mathcal{A}) + \sum_{\mu \in \mathbb{C}} d_j(\mu, \mathcal{A}) \right),$$

(1)

$$\dim X - \dim Y = r(\mathcal{A}) - l(\mathcal{A}).$$

(2)

The set of all those $\lambda \in \mathbb{C}$ for which $d_j(\lambda, \mathcal{A}) > 0$ for some $j > 0$ will be called the set of Jordan eigenvalues of $\mathcal{A}$ and denoted by $\sigma(\mathcal{A})$; the number $\sum_{k \geq 1} kd_k(\lambda, \mathcal{A})$ is called the multiplicity of $\lambda$ as the $J$-eigenvalue of $\mathcal{A}$.

Two $(X, Y)$ pencils $\mathcal{A}, \mathcal{B}$ are called equivalent if there exist two invertible operators $V: X \to X, W: Y \to Y$ such that $\mathcal{B}(\lambda) = W \mathcal{A}(\lambda) V$ for all $\lambda \in \mathbb{C}$. Two pencils are equivalent if and only if they have the same Kronecker structure [4].

KRONECKER STRUCTURE OF A PERTURBED PENCIL

Perturbation theorems for operators may be obtained with the help of functional analysis. There is no such calculus for operator pencils. We shall see that the tensor calculus may be useful in obtaining information about perturbed pencils.

Let $\mathcal{A}_n$ be a sequence of $(X, Y)$-pencils converging to $\mathcal{A}$ with $n \to \infty$, i.e.,

$$\mathcal{A}_n(\lambda) \to \mathcal{A}(\lambda) \quad \text{for all} \quad \lambda \in \mathbb{C}$$

(or equivalently for two different values of $\lambda$), and let $\mathcal{F}_n$ be a sequence of $(U, V)$-operator pencils converging to $\mathcal{F}$. Let us consider the operator $\langle \mathcal{F}, \mathcal{A} \rangle$ defined by

$$\langle \mathcal{F}, \mathcal{A} \rangle = [\mathcal{F}(\infty) \otimes \mathcal{A}(0)] - [\mathcal{F}(0) \otimes \mathcal{A}(\infty)].$$

It is clear that $\langle \mathcal{F}_n, \mathcal{A}_n \rangle$ converges to $\langle \mathcal{F}, \mathcal{A} \rangle$ and therefore $\text{null}(\langle \mathcal{F}_n, \mathcal{A}_n \rangle) \leq \text{null}(\langle \mathcal{F}, \mathcal{A} \rangle) = \text{the dimension of the kernel of } \langle \mathcal{F}, \mathcal{A} \rangle$, for sufficiently
large $n$. We shall study the relation between $\text{null}(\mathcal{F}, \mathcal{A})$ and the Kronecker structures of the pencils $\mathcal{F}, \mathcal{A}$; these relations will enable us to obtain some relations between the Kronecker structures of a pencil $\mathcal{A}$ and any pencil $\mathcal{B}$ sufficiently close to $\mathcal{A}$. Let $\mathcal{F} = \bigoplus \mathcal{F}_i$ and $\mathcal{A} = \bigoplus \mathcal{A}_j$ be decompositions of $\mathcal{F}$ and $\mathcal{A}$ into a direct sum of indecomposable pencils. Then

$$\langle \mathcal{F}, \mathcal{A} \rangle = \bigoplus_{i,j} \langle \mathcal{F}_i, \mathcal{A}_j \rangle,$$

and this implies that

$$\text{null}(\mathcal{F}, \mathcal{A}) = \sum_{i,j} \text{null}(\mathcal{F}_i, \mathcal{A}_j). \quad (3)$$

The nullities of $\langle \mathcal{F}, \mathcal{A} \rangle$ for all possible pairs $\mathcal{F}, \mathcal{A}$ of indecomposable pencils are evaluated in Lemma 1.

**Lemma 1.**

$$\text{null}(\mathcal{L}_j, \mathcal{L}_k) = 0,$$
$$\text{null}(\mathcal{L}_j, \mathcal{R}_k) = \text{null}(\mathcal{R}_k, \mathcal{L}_j) = (j - k)_+ \quad (= \max\{0, j - k\}),$$
$$\text{null}(\mathcal{R}_j, \mathcal{S}(\mu)) = \text{null}(\mathcal{S}(\mu), \mathcal{L}_j) = 0,$$
$$\text{null}(\mathcal{R}_j, \mathcal{R}_k) = j + k + 1,$$
$$\text{null}(\mathcal{R}_j, \mathcal{S}(\mu)) = \text{null}(\mathcal{S}(\mu), \mathcal{R}_j) = k,$$
$$\text{null}(\mathcal{S}(\mu), \mathcal{S}(\nu)) = \begin{cases} 0 & \text{if } \mu \neq \nu, \\ \min\{j, k\} & \text{if } \mu = \nu. \end{cases}$$

**Proof.** To prove the lemma completely one has to perform many elementary and simple calculations on matrices. As an example we shall prove only the last equality in the case when $\mu, \nu$ are finite numbers. The operator $\langle \mathcal{S}(\mu), \mathcal{S}(\nu) \rangle$ may be represented by the following $jk \times jk$ matrix:

$$\begin{bmatrix} (\nu - \mu)I_j - N_j & I_j \\ (\nu - \mu)I_j - N_j & \ddots & \ddots & \ddots \\ \ddots & \ddots & \ddots & I_j \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ \nu - \mu)I_j - N_j & \ddots & \ddots & \ddots & \ddots & I_j \\ (\nu - \mu)I_j - N_j & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots & \ddots & \ddots \\ & & & & & \ddots & \ddots & \ddots \\ & & & & & & \ddots & \ddots \\ & & & & & & & \ddots \\ & & & & & & & & \ddots \end{bmatrix}.$$
The determinant of this block bidiagonal matrix equals $(v - \mu)^k$, and is not 0 when $v \neq \mu$; thus clearly $\text{null}\langle \mathcal{H}(\mu), \mathcal{H}(v) \rangle = 0$. When $\mu = v$, then postmultiplying the above matrix by a nonsingular matrix

$$
\begin{bmatrix}
0 & 0 & \cdots & 0 & I_j \\
I_j & 0 & \cdots & 0 & N_j \\
N_j & I_j & \cdots & 0 & N_j^2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
N_j^{k-2} & N_j^{k-3} & \cdots & I_j & N_j^{k-1}
\end{bmatrix}
$$

and premultiplying the result by the nonsingular matrix

$$
\begin{bmatrix}
I_j \\
\cdot \\
\cdot \\
\cdot \\
N_j^{k-1} \\
\cdot \\
\cdot \\
\cdot \\
N_j & I_j
\end{bmatrix}
$$

we obtain the block-diagonal matrix $\text{diag}(I_j, \ldots, I_j, -N_j^k)$. Hence $\text{null}\langle \mathcal{H}(\mu), \mathcal{H}(\mu) \rangle = \text{null } N_j^k = \min\{j, k\}$. $\blacksquare$

The following corollary follows easily from Lemma 1 and the identities (3), (1), (2).

**COROLLARY 1.** For any $(X, Y)$-pencil $\mathfrak{A}$ the following identities hold:

$$
\text{null}\langle \mathcal{B}_j, \mathfrak{A} \rangle = \dim X + j(\dim X - \dim Y) + \sum_k (j - k) \cdot l_k(\mathfrak{A}),
$$

$$
\text{null}\langle \mathcal{L}_j, \mathfrak{A} \rangle = \sum_k (j - k) \cdot r_k(\mathfrak{A}),
$$

$$
\text{null}\langle \mathcal{H}(\mu), \mathfrak{A} \rangle = j r(\mathfrak{A}) + \sum_k \min\{j, k\} \cdot d_k(\mu, \mathfrak{A}).
$$

Now we can state and prove two general theorems on the Kronecker structure of a perturbed pencil.
THEOREM 1. Let a sequence \( \mathcal{A}_n \) of \((X,Y)\) operator pencils converge to a pencil \( \mathcal{B} \), and let \( \sigma_\varepsilon \) be the \( \varepsilon \)-neighborhood of \( \sigma_\varepsilon (\mathcal{B}) \). Then for sufficiently large \( n \) the following inequalities hold (simultaneously):

\[
\sum_k (j-k) [r_k(\mathcal{B}) - r_k(\mathcal{A}_n)] \geq 0, \tag{4}
\]

\[
\sum_k (j-k) [l_k(\mathcal{B}) - l_k(\mathcal{A}_n)] \geq 0, \tag{5}
\]

\[
j [r(\mathcal{B}) - r(\mathcal{A}_n)] + \sum_k \min \{j,k\} [d_k(\mu, \mathcal{B}) - d_k(\mu, \mathcal{A}_n)] \geq 0 \tag{6}
\]

for all \( j = 1, 2, 3, \ldots \) and \( \mu \in (\overline{\mathbb{C}} \setminus \sigma_\varepsilon) \cup \sigma_\varepsilon (\mathcal{B}) \).

Proof. Since a finite-dimensional space has only one natural topology, we may assume without loss of generality that all vector spaces under consideration are Hilbert spaces. We shall show firstly that for a fixed \( j \) and sufficiently large \( n \), (6) holds for all \( \mu \in K_1 \setminus \sigma_\varepsilon \), where \( K_1 = \{ z \in \mathbb{C}; |z| \leq 1 \} \). It follows from Corollary 1 that for \( \mu \in K_1 \setminus \sigma_\varepsilon \) we have \( \text{null} \langle \mathcal{I}_j(\mu), \mathcal{B} \rangle = jr(\mathcal{B}) \). This and the fact that the operator function \( \langle \mathcal{I}_j(\mu), \mathcal{B} \rangle \) is a continuous function of \( \mu \) imply that the kernel of \( \langle \mathcal{I}_j(\mu), \mathcal{B} \rangle \) depends continuously on \( \mu \in K_1 \setminus \sigma_\varepsilon \). Since the set \( K_1 \setminus \sigma_\varepsilon \) is compact, there exists a positive number \( c \) such that for all \( \mu \in K_1 \setminus \sigma_\varepsilon \), \( ||\langle \mathcal{I}_j(\mu), \mathcal{B} \rangle x|| \geq c ||x|| \) for all \( x \) orthogonal to \( \ker \langle \mathcal{I}_j(\mu), \mathcal{B} \rangle \). When \( n \) is so large that the inequality

\[
||\langle \mathcal{I}_j(\mu), \mathcal{B} \rangle - \langle \mathcal{I}_j(\mu), \mathcal{A}_n \rangle x|| \leq \frac{c}{2} ||x||
\]

holds for all vectors \( x \) and all \( \mu \in K_1 \setminus \sigma_\varepsilon \), then for all such \( n \) we have \( ||\langle \mathcal{I}_j(\mu), \mathcal{A}_n \rangle x|| \geq (c/2)||x|| \) for all \( x \) orthogonal to \( \ker \langle \mathcal{I}_j(\mu), \mathcal{B} \rangle \). This implies that for sufficiently large \( n \)

\[
\text{null} \langle \mathcal{I}_j(\mu), \mathcal{A}_n \rangle \leq \text{null} \langle \mathcal{I}_j(\mu), \mathcal{B} \rangle \quad \text{for all} \quad \mu \in K_1 \setminus \sigma_\varepsilon.
\]

Repeating the above argument for the continuous operator function \( \mu^{-1} \langle \mathcal{I}_j(\mu), \mathcal{B} \rangle \) in the compact domain \( \overline{\mathbb{C}} \setminus \sigma_\varepsilon \setminus \{ z \in \mathbb{C}; |z| < 1 \} \), we show that for almost all \( n \)

\[
\text{null} \langle \mathcal{I}_j(\mu), \mathcal{A}_n \rangle \leq \text{null} \langle \mathcal{I}_j(\mu), \mathcal{B} \rangle
\]
for all \( \mu \) from this domain, and therefore also for all \( \mu \in \mathbb{C} \setminus \sigma_e \). This and similar considerations show that for sufficiently large \( n \), all \( j \leq \max\{ \dim X, \dim Y \} - m \), and all \( \mu \in \sigma_f(\mathcal{B}) \cup (\mathbb{C} \setminus \sigma_f) \), the following inequalities hold simultaneously:

\[
\text{null}\langle \mathcal{L}_j, \mathcal{A}_n \rangle \leq \text{null}\langle \mathcal{L}_j, \mathcal{B} \rangle,
\]

\[
\text{null}\langle \mathcal{R}_j, \mathcal{A}_n \rangle \leq \text{null}\langle \mathcal{R}_j, \mathcal{B} \rangle,
\]

\[
\text{null}\langle \mathcal{I}_j(\mu), \mathcal{A}_n \rangle \leq \text{null}\langle \mathcal{I}_j(\mu), \mathcal{B} \rangle.
\]

Evaluating the nullities of the operators involved with help of Corollary 1, we see that for all the above values of \( n \), \( j \), and \( \mu \) the inequalities (4), (5), (6) hold.

Now note that for \( k > m \) we always \( r_k(\mathcal{A}_n) = l_k(\mathcal{A}_n) = d_k(\mu, \mathcal{A}_n) = 0 \), and that in view of (2), \( l(\mathcal{B}) - l(\mathcal{A}_n) = r(\mathcal{B}) - r(\mathcal{A}_n) \). Therefore the left-hand sides of (4), (5), (6) for \( j = m + p \) differ from the left-hand sides of these inequalities (respectively) for \( j = m \) on \( p[r(\mathcal{A}_n) - r(\mathcal{B})] \). The inequality (6) for \( j = 1 \) and some \( \mu \in \sigma_f(\mathcal{A}_n) \cup \sigma_f(\mathcal{B}) \) takes the form \( r(\mathcal{A}_n) \leq r(\mathcal{B}) \). These facts imply that when the inequalities (4), (5), (6) hold for \( j \leq m \), then these inequalities hold also for all \( j > m \). This ends the proof.

The next perturbation theorem describes the behavior of Jordan eigenvalues of pencils \( \mathcal{A}_n \) approximating a pencil \( \mathcal{B} \) in the case when all these pencils have the same number of blocks of type \( \mathcal{R} \) in their canonical Kronecker decomposition, and hence also [cf. (2)] the same number of blocks of type \( \mathcal{L} \). This assumption is equivalent to \( \text{null } \sigma_f(\mathcal{X}) = \text{null } \sigma_f(\mathcal{B}) \) for all \( \mu \in \mathbb{C} \setminus [\sigma_f(\mathcal{A}_n) \cup \sigma_f(\mathcal{B})] \).

**THEOREM 2.** Suppose that \((X, Y)\)-pencils \( \mathcal{A}_n \) converge to a pencil \( \mathcal{B} \) and that \( r(\mathcal{A}_n) = r(\mathcal{B}) \), \( n = 1, 2, 3, \ldots \). Then

(i) for any \( \epsilon \)-neighborhood \( \sigma_e \) of \( \sigma_f(\mathcal{B}) \)

\[ \sigma_f(\mathcal{A}_n) \subset \sigma_e \quad \text{for sufficiently large } n, \]

(ii) if \( \mu \in \sigma_f(\mathcal{B}) \) and \( \Omega \) is a closed neighborhood of \( \mu \) (in \( \mathbb{C} \)) such that \( \Omega \cap \sigma_f(\mathcal{B}) = \{ \mu \} \), then the total multiplicity of \( J \)-eigenvalues of \( \mathcal{A}_n \) contained in \( \Omega \) is not greater than the multiplicity of \( \mu \) as a \( J \)-eigenvalue of \( \mathcal{B} \) for sufficiently large \( n \), i.e.,

\[
\sum_{\nu \in \Omega} \sum_k k d_k(\nu, \mathcal{A}_n) \leq \sum_k k d_k(\mu, \mathcal{B}). \tag{7}
\]
Proof. (i) is a direct consequence of the inequality (6) from Theorem 1.

In order to prove the second assertion we may assume additionally (without loss of generality) that the total multiplicity of \( J \)-eigenvalues of \( \mathscr{A}_n \) contained in \( \Omega \) is constant. Thus let \( \lambda_{1,n}, \lambda_{2,n}, \ldots, \lambda_{s,n} \) be all \( J \)-eigenvalues of \( \mathscr{A}_n \) contained in \( \Omega \) and repeated according to multiplicity. Let \( \mu_{i,n} (1 \leq i \leq r_n) \) be all different \( J \)-eigenvalues of \( \mathscr{A}_n \) contained in \( \Omega \), and let \( k_{i,n} \) be the multiplicity of \( \mu_{i,n} \). Passing to a subsequence, we may assume that also the numbers \( r_n, k_{i,n} \) do not depend on \( n \), i.e., \( r_n = r, k_{i,n} = k_i \).

We define a sequence of \( (\mathbb{C}^s, \mathbb{C}^s) \)-pencils \( \mathcal{F}_n \) convergent to the pencil \( \mathcal{F}_s(\mu) \):

\[
\mathcal{F}_n(\lambda) = \begin{bmatrix}
\lambda - \lambda_{1,n} & 1 \\
\lambda - \lambda_{2,n} & \ddots & \ddots \\
& \ddots & \ddots & 1 \\
& & \ddots & \lambda - \lambda_{s,n}
\end{bmatrix}
\text{ if } \mu \in \mathbb{C},
\]

\[
\mathcal{F}_n(\lambda) = \begin{bmatrix}
1 - \lambda \cdot \lambda_{1,n} & \lambda \\
1 - \lambda \cdot \lambda_{2,n} & \ddots & \ddots \\
& \ddots & \ddots & \lambda \\
& & \ddots & 1 - \lambda \cdot \lambda_{s,n}
\end{bmatrix}
\text{ if } \mu = \infty
\]

(with the convention \( \infty^{-1} = 0 \)). Note that the pencil \( \mathcal{F}_n \) is equivalent to \( \otimes_{i=1}^s \mathcal{I}_{k_i}(\mu_{i,n}) \). This follows from the facts that

\[
\text{null } \mathcal{F}_n(\lambda) = \begin{cases}
0 & \text{if } \lambda \neq \lambda_{i,n} \text{ for } i = 1, 2, \ldots, s, \\
1 & \text{if } \lambda = \lambda_{i,n} \text{ for some } i,
\end{cases}
\]

and that \( \det \mathcal{F}_n(\lambda) = \prod_{i=1}^s (\lambda - \lambda_{i,n}) = \prod_{i=1}^s (\lambda - \mu_{i,n})^{k_i} \) if \( \mu \in \mathbb{C} \), and \( \det \mathcal{F}_n(\lambda) = \prod_{i=1}^s (1 - \lambda \cdot \mu_{i,n})^{k_i} \) if \( \mu = \infty \). Since \( \langle \mathcal{F}_n, \mathscr{A}_n \rangle \to \langle \mathcal{I}_s(\mu), \mathcal{B} \rangle \) as \( n \to \infty \), we have \( \text{null } \langle \mathcal{F}_n, \mathscr{A}_n \rangle \leq \text{null } \langle \mathcal{I}_s(\mu), \mathcal{B} \rangle \) for \( n \) sufficiently large. We evaluate these nullities using (3) and Corollary 1:

\[
\text{null } \langle \mathcal{F}_n, \mathscr{A}_n \rangle = \sum_{i=1}^r \text{null } \langle \mathcal{I}_{k_i}(\mu_{i,n}), \mathscr{A}_n \rangle
- \sum_{i=1}^r \left( k_i r(\mathscr{A}_n) + \sum_k \min \{ k, k_i \} d_k(\mu_{i,n}, \mathscr{A}_n) \right).
\]
Since $k_i = \sum_k k d_k(\mu_{i,n}, \mathcal{A}_n) \geq k$ if $d_k(\mu_{i,n}, \mathcal{A}_n) > 0$, and $\sum_{i=1}^r k_i = s$, we have

$$\text{null}(\mathcal{F}_n, \mathcal{A}_n) = sr(\mathcal{B}) + s$$

$$\leq \text{null}(\mathcal{F}(\mu), \mathcal{B}) = sr(\mathcal{B}) + \sum_k \min\{s, k\} d_k(\mu, \mathcal{B})$$

$$\leq sr(\mathcal{B}) + \sum_k kd_k(\mu, \mathcal{B}).$$

This inequality implies (7), since $s = \sum_{\mu \in \mathcal{B}} k d_k(\nu, \mathcal{A}_n)$.

**Remark.** If $r(\mathcal{B}) = 0$ [or $l(\mathcal{B}) = 0$], then it follows from Theorem 1 [and (2)] that the assumption $r(\mathcal{A}_n) = r(\mathcal{B})$ of Theorem 2 is satisfied for sufficiently large $n$. In particular, if $\mathcal{B}$ is a regular pencil [i.e. $r(\mathcal{B}) = l(\mathcal{B}) = 0$], then the pencils $\mathcal{A}_n$ are also regular for large $n$, and then we have in fact the equality in (7).

One may ask whether these perturbation theorems may be improved. In the next section we shall show that neglecting small perturbations of $f$-eigenvalues of approximating pencils $\mathcal{A}_n$, not much more can be proved about their Kronecker structure.

More precisely: Suppose that a sequence $\mathcal{A}_n$ of $(X, Y)$-pencils converges to a pencil $\mathcal{B}$. Passing to a subsequence, we may assume that the numbers $r_j(\mathcal{A}_n), l_j(\mathcal{A}_n)$ do not depend on $n$. Then the pencils $\mathcal{A}_n$ have a constant number of $f$-eigenvalues (counting multiplicities), and therefore (since $C$ is compact) we may assume moreover that these $f$-eigenvalues converge to $\mu_1, \mu_2, \ldots, \mu_s$. For each $(X, Y)$-pencil $\mathcal{A}$ there are orthonormal bases of $X$ and $Y$ in which the pencil is represented in a quasitriangular form (cf. [7, Proposition 4.7])

$$\mathcal{A}(\lambda) = \begin{bmatrix} R(\lambda) & * & * \\ 0 & I(\lambda) & * \\ 0 & 0 & S(\lambda) \end{bmatrix}$$

such that

(i) $\mathcal{A}$ is equivalent to $R \oplus I \oplus S$,

(ii) $r_j(\mathcal{B}) = r_j(\mathcal{A}), \quad l_j(\mathcal{B}) = l_j(\mathcal{A}), \quad d_j(\mu, \mathcal{A}) = d_j(\mu, \mathcal{B})$ for $j = 0, 1, 2, \ldots, \mu \in \mathcal{C}$, and all other numbers in the Kronecker structures of these pencils are 0.

(iii) $I(\lambda)$ is an upper triangular matrix with elements $\alpha_i - \beta_i \lambda$, where $\alpha_i/\beta_i = \mu_i$ are $f$-eigenvalues of $I$ (and $\mathcal{A}$), with convention $\alpha/0 = \infty$; (cf. [6, Theorem 3.1].
When the pencils $\mathcal{A}_n$ converge, then the respective numbers $\alpha_{i,n}, \beta_{i,n}$ are uniformly bounded; therefore (passing once more to a subsequence) we may assume that $\alpha_{i,n} \to \alpha_i$, $\beta_{i,n} \to \beta_i$. Now, in the quasitriangular representation of $\mathcal{A}_n$, we replace the numbers $\alpha_{i,n}, \beta_{i,n}$ by: (i) $\alpha_i, \beta_i$ if $(\alpha_i, \beta_i) \neq (0,0)$; (ii) $(1/n)\mu_i, 1/n$ if $(\alpha_i, \beta_i) = (0,0)$ and $\mu_i \in \mathbb{C}$; (iii) $1/n, 0$ if $(\alpha_i, \beta_i) = (0,0)$ and $\mu_i = \infty$. The Kronecker structure of the pencils $\mathcal{A}_n$ obtained in this way differs slightly from that of $\mathcal{A}_n$. Moreover the sequence $\mathcal{A}_n$ converges to the same limit and contains a subsequence of equivalent pencils. In the next section we shall describe the Kronecker structures of all such possible subsequences.

**CLOSURE OF THE EQUIVALENCE ORBIT OF A MATRIX PENCIL**

Let $\mathcal{S}(\mathcal{A})$ (the equivalence orbit of a pencil $\mathcal{A}$) be the set of all pencils equivalent with $\mathcal{A}$. Then $\mathcal{S}(\mathcal{A})^-$, the closure of $\mathcal{S}(\mathcal{A})$, is the set of all pencils $\mathcal{B}$ for which there exist a sequence $\{\mathcal{A}_n\}^\infty_{n=1}$ of pencils equivalent to $\mathcal{A}$ such that $\mathcal{A}_n \to \mathcal{B}$. In order to characterize $\mathcal{S}(\mathcal{A})^-$ it suffices to find the relations between the Kronecker structure of $\mathcal{A}$ and that of $\mathcal{B}$, for all $\mathcal{B} \in \mathcal{S}(\mathcal{A})^-$. Necessary conditions for $\mathcal{B} \in \mathcal{S}(\mathcal{A})^-$ may be easily obtained from Theorem 1; they constitute the "if" part of Theorem 3.

**THEOREM 3.** A pencil $\mathcal{B}$ lies in the closure of the equivalence orbit of a pencil $\mathcal{A}$ if and only if the following inequalities hold:

$$\sum_{k}(j - k), r_k(\mathcal{A}) \leq \sum_{k}(j - k), r_k(\mathcal{B}),$$

$$\sum_{k}(j - k), l_k(\mathcal{A}) \leq \sum_{k}(j - k), l_k(\mathcal{B}),$$

$$jr(\mathcal{A}) + \sum_{k}\min\{j, k\} d_k(\mu, \mathcal{A}) \leq jr(\mathcal{B}) + \sum_{k}\min\{j, k\} d_k(\mu, \mathcal{B})$$

for all $j = 1, 2, 3, \ldots$ and $\mu \in \mathbb{C}$.

To show that the necessary conditions for $\mathcal{B} \in \mathcal{S}(\mathcal{A})$ given in Theorem 3 are also sufficient we introduce partial orders on $(X,Y)$-pencils. We shall write $A \preceq B$ if for two $(X,Y)$-pencils $A$, $B$ the inequalities (8), (9), (10) hold for all values of $j$ and $\mu$. We shall write $A \succeq B$ if $B \in \mathcal{S}(\mathcal{A})^-$. The strict relation $A \prec B$ ($A \succeq B$) denotes that $A \preceq B$ ($A \succeq B$) but the inverse relation does not hold. Theorem 3 may be equivalently formulated in the following way: the relations $\preceq$, $\succeq$ are equivalent. Simple properties of these relations are collected in Lemmata 2 and 3.
LEMMA 2.

(i) The relations $\prec$ and $\preceq$ are transitive.
(ii) $A \preceq B$ implies $A \prec B$.
(iii) The relations $A \preceq B$ and $B \preceq A$ imply that the pencils $A, B$ are equivalent.
(iv) If $A \preceq B$ then $A \oplus C \preceq B \oplus C$ for any pencil $C$.

Proof. Everything but the transitivity of $\preceq$ is clear. Let $\| \cdot \|$ be an operator norm. If $\| z \| = \| z' \|$ and $\| z'' \| = \| z''' \|$, then for any positive number $\varepsilon$ there exist invertible operators $W, W_1, V, V_1$ such that $\| g(y) - W_2 g(y) V \| + \| g(y) - W_1 g(y) V_1 \| \leq \varepsilon/2$ and $\| g(y) - W_2 g(y) V_1 \| + \| g(\infty) - W_1 g(\infty) V_1 \| \leq \varepsilon/2 \| W \||V\|$. Then $WW_1, VV_1$ are invertible operators, and it is easy to verify that $\| g(y) - WW_1 g(0) V_1 V \| \leq \varepsilon$ and $\| g(\infty) - WW_1 g(\infty) V_1 V \| \leq \varepsilon$. This shows that $A \preceq C$.

LEMMA 3. There exists no infinite chain $A_1 < A_2 < \cdots$ of $(X, Y)$-pencils.

Proof. Assume the contrary. In view of (1) and (2), there are only a finite number of sequences $\{ l_j(A_n) \}_{n=0}^\infty, \{ r_j(A_n) \}_{n=0}^\infty$ which correspond to $(X, Y)$-pencils. Therefore, passing to a subsequence we may assume that for a strictly ordered sequence $\{ z_j \}_{j=0}^\infty$ the numbers $l_j(A_n), r_j(A_n)$ do not depend on $n$. For $j > \dim X$ the inequality (10) takes now the form $\sum k d_k(\lambda, A_n) \leq \sum k d_k(\lambda, A_{n+1})$, and it follows from (1) that the sum $\sum_{\lambda \in C} \sum_k k d_k(\lambda, A_n)$ does not depend on $n$. The assertion of the lemma follows from the fact that for each $\lambda$ there are only a finite number of points $\lambda \in C$ for which $\sum_{k=1}^\infty \sum_{\lambda \in C} k [d_k(\lambda, A_n)] = 0$ in nonnegative integers $d_1, d_2, \ldots$.

Examples of consecutive pencils $A < B$ (i.e. such that there is no $C$ such that $A < C < B$) are given in Lemma 5. We shall not prove that these pencils are really consecutive; it is easy to verify. Moreover, investigating the proof of Lemma 6, one can see that if $A < B$ are two consecutive pencils, then $A = A_1 \oplus A_2, B = B_1 \oplus B_2$, where the pencils $A_1, B_1$ are equivalent, while the pencils $A_2, B_2$ are equivalent to one of the ordered pairs of pencils presented in Lemma 5.
The following technical result will be used in the proof of Lemma 5.

**Lemma 4.** Let \( x(\lambda) \) be a nonzero vector such that \( R_k(\lambda)x(\lambda) = 0 \) and let \( \lambda_0, \lambda_1, \lambda_2, \ldots \) be distinct points of \( \overline{C} \). Then the vectors \( x(\lambda_0), x(\lambda_1), \ldots, x(\lambda_r) \) are linearly independent if and only if \( r \leq k \).

**Proof.** It suffices to note that when \( r = k \), the vectors \( x(\lambda_i) (0 \leq i \leq k) \) are (up to multiplicative constants) the columns of a nonsingular Vandermonde matrix \( \{\lambda_j^i\}_{i,j=0}^{k} \). In the case when \( \lambda_0 = \infty \) the first column should be replaced by \((0,0,\ldots,0,1)^T\).

**Lemma 5.** The following relations hold:

(i) \( R_j \oplus R_k < R_{j+1} + R_{k+1}, 1 \leq j \leq k, \)

(ii) \( L_j \oplus L_k < L_{j+1} \oplus L_{k+1}, 1 \leq j \leq k, \)

(iii) \( R_j \oplus I_{k+1}(\mu) < R_j \oplus I_{k+1}(\mu), j, k = 0, 1, 2, \ldots, \mu \in \overline{C}, \)

(iv) \( L_j \oplus I_{k+1}(\mu) < L_j \oplus I_{k+1}(\mu), j, k = 0, 1, 2, \ldots, \mu \in \overline{C}, \)

(v) \( I_j \oplus I_{k+1}(\mu) < I_j \oplus I_{k+1}(\mu), 1 \leq j \leq k, \mu \in \overline{C}, \)

(vi) \( \bigoplus I_{k+1}(\lambda_i) \cong L_p \oplus R_q \) if \( \lambda_i \neq \lambda_j \) for \( i \neq j, \lambda_i, \lambda_j \in \overline{C}, p + q + 1 = \Sigma_{i=1}^s k_i. \)

(\( \mathcal{I}_\emptyset(\cdot) \) denotes the \((\{0\}, \{0\})\)-pencil.)

**Proof.** (i): Assume \( s, t \in (0,1) \), and consider the following \((j + k) \times (j + k + 2)\) matrix pencil:

\[
\mathcal{I}_{s,t}(\lambda) = \begin{bmatrix}
-\lambda & 1 & & & & \\
-\lambda & 1 & & & & \\
& \ddots & \ddots & \ddots & \ddots & \\
& & -\lambda & 1 & & \\
& & -s\lambda & s-t\lambda & t & \\
& & & -\lambda & 1 & \\
& & & & \ddots & \\
& & & & & -\lambda & 1
\end{bmatrix}_{jth \ row.}
\]
For any \( \lambda \in \mathbb{C} \) all the rows of the matrix \( \mathcal{A}_{s,t}(\lambda) \) are linearly independent and null \( \mathcal{A}_{s,t}(\lambda) = 2 \). This with (2) and (6), for \( j = 1 \), implies that \( \mathcal{A}_{s,t} \) is equivalent to \( \mathcal{R}_p \oplus \mathcal{R}_r \), where \( p, r \) are nonnegative integers such that \( p + r = j + k \), \( p \leq r \).

Let \( x(\lambda), y(\lambda) \) be linearly independent vectors such that \( \mathcal{A}_{s,t}(\lambda)x(\lambda) = \mathcal{A}_{s,t}(\lambda)y(\lambda) = 0 \), and let \( \lambda_0 = s/t \), \( \lambda_1, \lambda_2, \ldots \) be different complex numbers. The vectors \( x(\lambda), y(\lambda) \) may be chosen in the following way:

\[
x(\lambda_0) = (0, \ldots, 0, 1, 0, \ldots, 0)^T,
\]

\[
y(\lambda_0) = (1, \lambda_0, \lambda_0^2, \ldots, \lambda_0^{j+k+1})^T;
\]

\[
x(\lambda) = \left( 1, \lambda, \lambda^2, \ldots, \lambda^{j-1}, s\lambda^i(s - t\lambda)^{-1}, 0, \ldots, 0 \right)^T,
\]

\[
y(\lambda) = \left( 0, \ldots, 0, -t(s - t\lambda)^{-1}, 1, \lambda, \lambda^2, \ldots, \lambda^k \right)^T \quad \text{for} \quad \lambda \neq \lambda_0.
\]

It follows from Lemma 4 that the vectors \( x(\lambda_i), y(\lambda_i), 0 \leq i \leq q \), are linearly independent if and only if \( q \leq p \). Since the vectors \( x(\lambda_i), 0 \leq i \leq j + 1 \), are linearly dependent, we have \( p \leq j \). To show that \( p = j \) it suffices to verify that the vectors \( x(\lambda_i), y(\lambda_i), 0 \leq i \leq j \), are linearly independent, and this follows from the fact that the matrix

\[
\begin{bmatrix}
1 & \cdots & 1 \\
\lambda_1 & \cdots & \lambda_j \\
\vdots & & \vdots \\
\lambda_1^{j-1} & \cdots & \lambda_j^{j-1} \\
1 & \cdots & 1
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 \\
\lambda_0 \\
\vdots \\
\lambda_0^{j-1} \\
1 & \cdots & 1
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & \cdots & 1 \\
\lambda_1 & \cdots & \lambda_j \\
\vdots & & \vdots \\
\lambda_1^{j+1} & \cdots & \lambda_j^{j+1}
\end{bmatrix}
\]

—whose columns consist of the first \( 2j + 2 \) components of the vectors

\[
x(\lambda_i) - s\lambda_i^j(s - t\lambda_i)^{-1}x(\lambda_0), \quad i = 1, 2, \ldots, j. \quad x(\lambda_0). \quad y(\lambda_i) + t(s - t\lambda_i)^{-1}x(\lambda_0), \quad i = 1, 2, \ldots, j, \quad y(\lambda_0)— \text{is nonsingular.}
\]
In this way we have shown that $\mathcal{A}_{s, t}$ is equivalent to $\mathcal{B}_j \oplus \mathcal{B}_k$. Claim (i) follows from the convergence $\mathcal{A}_{s, 1-s} \to \mathcal{B}_j \oplus \mathcal{B}_{k+1}$ when $s \searrow 0$.

(ii): Considering dual (transposed) pencils, we obtain claim (ii) from claim (i).

(iii): It suffices to prove claim (iii) for $\mu = 0$ only, since the pencil $\mathcal{B}_j(\lambda) \oplus \mathcal{F}_k(\nu, \lambda)$ is equivalent to the pencil

$$\left[ \lambda \mathcal{B}_j(\infty) + \{ \mathcal{B}_j(0) - \nu \mathcal{B}_j(\infty) \} \right] \oplus \left[ \lambda \mathcal{F}_k(0, \infty) + \{ \mathcal{F}_k(0, 0) - \nu \mathcal{F}_k(0, \infty) \} \right]$$

when $\nu \in \mathbb{C}$, and to $[\lambda \mathcal{B}_j(0) + \mathcal{B}_j(\infty)] \oplus [\lambda \mathcal{F}_k(0, 0) + \mathcal{F}_k(0, \infty)]$ when $\nu = \infty$.

Let us consider the following $(j+k+1) \times (j+k+2)$ matrix pencil [for $s, t \in (0, 1)$]:

$$\mathcal{B}_{s, t}(\lambda) =$$

$$\begin{bmatrix}
-\lambda & 1 & & & & \\
-\lambda & -\lambda & 1 & & & \\
& & & & & \\
& & & & & \\
& & & & & \\
& & & & & \\
& & & & & \\
\end{bmatrix}$$

(j+1)th row

It is easy to check that null $\mathcal{B}_{s, t}(\lambda) = 1$ for all $\lambda \in \overline{\mathbb{C}} \setminus \{0\}$, and null $\mathcal{B}_{s, t}(0) = 2$. This implies that $\mathcal{B}_{s, t}$ is equivalent to $\mathcal{B}_j \oplus \mathcal{F}_{k+j+1-r}(0)$ for some integer $r$.

If $\lambda_0 = s/t, \lambda_1, \lambda_2, \ldots$ are different nonzero complex numbers, then

$$x_0 = (0, \ldots, 0, 1, 0, \ldots, 0)^T \in \ker \mathcal{B}_{s, t}(\lambda_0),$$

$$x_i = (1, \lambda_i, \lambda_i^2, \ldots, \lambda_i^j, s \lambda_i^{j+1}(s - t \lambda)^{-1}, 0, \ldots, 0)^T \in \ker \mathcal{B}_{s, t}(\lambda_i).$$

Since the vectors $x_0, x_1, \ldots, x_p$ are linearly independent if and only if $p \leq j + 1$, Lemma 4 implies that $r = j + 1$, i.e., $\mathcal{B}_{s, t} \in S(\mathcal{B}_j \oplus \mathcal{F}_k(0))$.

Claim (iii) now follows from the convergence $\mathcal{B}_{s, 1-s} \to \mathcal{B}_j \oplus \mathcal{F}_k(0)$ when $s \to 0$.

(iv): This claim is a consequence of (iii), when transposed matrices are considered.
(v): As in case (iii), it suffices to consider only the case \( \mu = 0 \). We define the following \((k + j)\)-square matrix pencil:

\[
E_{s,t}(\lambda) = \begin{bmatrix}
\lambda & 1 & \cdots & \cdots & \cdots & s \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
& & \ddots & \ddots & \ddots & t \\
\vdots & & & \ddots & \ddots & \lambda \\
& & & & \ddots & \lambda \\
& & & & & 1 \\
\end{bmatrix}
\]

\[= \lambda I_{j+k} + \begin{bmatrix}
N_{j-1} & B_s \\
0 & C_t
\end{bmatrix}.
\]

The matrix \( E_{s,t}(\lambda) \) is nonsingular for all \( \lambda \in \mathbb{C} \setminus \{0\} \), and null \( E_{s,t}(0) = 2 \), if \( s, t \neq 0 \). Therefore \( E_{s,t} \) is equivalent to \( \mathcal{J}_p(0) \oplus \mathcal{J}_{k+j-p}(0) \), where \( p \) is some integer. It is easy to verify that \( B_s C_t = 0 \), and further that the matrix

\[
[E_{s,t}(0)]^n = \begin{bmatrix}
N_{j-1}^n & N_{j-1}^{n-1}B_s \\
0 & C_t^n
\end{bmatrix}
\]

is nonzero for \( n < k + 1 \) and zero for \( n = k + 1 \). Therefore \( E_{s,t} \) is equivalent to \( \mathcal{J}_{j-1}(0) \oplus \mathcal{J}_{k+1}(0) \) for nonzero numbers \( s, t \). Claim (v) follows from the fact that when \( s \neq 1 \) the pencil \( E_{s,1-s} \) converges to the pencil \( \mathcal{J}_j(0) \oplus \mathcal{J}_{k}(0) \).

(vi): Let \( \mu_j = \lambda \), when \( \sum_{a=1}^{i-1} k_a \leq j \leq \sum_{a=1}^{i} k_a \), and note that for \( \epsilon > 0 \) the pencil

\[
\mathcal{D}_{\epsilon}(\lambda) = \begin{bmatrix}
\lambda - \mu_0 & 1 \\
\lambda - \mu_1 & \ddots & \ddots \\
& \ddots & \ddots & \ddots & \ddots \\
& & \ddots & \ddots & 1 \\
& & & \epsilon(\lambda - \mu_{q} - 1) & 1 \\
& & & & \lambda - \mu_{q+1} & \ddots \\
& & & & & \ddots & 1 \\
& & & & & & \lambda - \mu_{p+q}
\end{bmatrix}
\]
is equivalent to $\bigoplus_{i=1}^{s} F_i(\lambda_i)$. This follows from the facts that

$$\text{null } D_\lambda(\lambda) = \begin{cases} 0 & \text{if } \lambda \neq \lambda_i \text{ for } i = 1, 2, \ldots, s, \\
1 & \text{if } \lambda = \lambda_i \text{ for some } i \ (1 \leq i \leq s), \end{cases}$$

and that $\det D_\lambda(\lambda) = \epsilon \prod_{i=1}^{s}(\lambda - \lambda_i)^{h_i}$.

To finish the proof it suffices to show that $D_0$ is equivalent to $L_0 \oplus L_q$. Since $\text{null } D_0(\lambda) = 1$ for all $\lambda \in \mathbb{C}$, $D_0$ is similar to $L_0 \oplus L_k$, where $j + k = p + q$. The vector $x(\lambda) = (1, \lambda - p \mu, (\lambda - p \mu)(\lambda - q \mu), \ldots, \lambda \mu^{j-1})^T$ belongs to the kernel of $D_0(\lambda)$; $D_0(\lambda)x(\lambda) = 0$. If $v_i$, $i = 0, 1, 2, \ldots$, are different complex numbers, then the vectors $x(v_i)$, $i = 0, 1, \ldots, r$, are linearly independent if and only if $r \leq q$; this follows from the fact that the polynomials in $\lambda$, which are the consecutive coordinates of $x(\lambda)$, are then linearly independent. Hence by Lemma 4, $k = q$; this means that $D_0$ is equivalent to $L_0 \oplus L_q$.

**Lemma 6.** If $\mathcal{A} < \mathcal{B}$, then there exists a pencil $\mathcal{C}$ such that

$$\mathcal{A} < \mathcal{C} \leq \mathcal{B} \text{ or } \mathcal{A} \leq \mathcal{C} < \mathcal{B}.$$ 

**Proof.** Note that it suffices to prove the lemma in the case when the pencils $\mathcal{A}, \mathcal{B}$ have no common elementary pencils in their canonical Kronecker decompositions, i.e.

$$l_j(\mathcal{A})l_j(\mathcal{B}) - r_j(\mathcal{A})r_j(\mathcal{B}) = d_j(\lambda, \mathcal{A})d_j(\lambda, \mathcal{B}) = 0$$

for all $j = 0, 1, 2, \ldots$ and $\lambda \in \overline{\mathbb{C}}$; (11)

this is assumed in the sequel.

The general idea of the proof is to decompose a pencil $\mathcal{A}$ into a direct sum $\mathcal{A} = \mathcal{A}_0 \oplus \mathcal{A}_1$ in such a way that $\mathcal{A}_1$ is equivalent to one of the pencils considered on Lemma 5, then to define $\mathcal{A}_2$ as a pencil standing on the other side of a respective relation from Lemma 5, and then to verify that the pencil $\mathcal{C} = \mathcal{A}_0 \oplus \mathcal{A}_2$ satisfies $\mathcal{A} \leq \mathcal{C} \leq \mathcal{B}$. To realize this idea we consider eight special cases. Each pair $\mathcal{A}, \mathcal{B}$ of ordered pencils enters into at least one of these cases.

(i) $l(\mathcal{A}) \geq 2$. Let $r$ be the smallest natural number such that $l_r(\mathcal{A}) \neq 0$, and $s$ the smallest natural number such that $\sum_{j=r}^{r+s} l_j(\mathcal{A}) \geq 2$. The assumption
(11) implies that there exists $t$ such that $0 \leq t < r$ and $l_t(\mathcal{B}) \neq 0$. The pencil $\mathcal{A}$ may be represented as a direct sum of pencils $\mathcal{A} = \mathcal{A}_0 \oplus \mathcal{L}_r \oplus \mathcal{L}_s$. We put $\mathcal{C} = \mathcal{A}_0 \oplus \mathcal{L}_{r-1} \oplus \mathcal{L}_{s+1}$. By Lemma 5 we have $\mathcal{A} \prec \mathcal{C}$.

Since $l_k(\mathcal{C}) = l_k(\mathcal{A}) + \delta_{k-r-1} - \delta_{k-r} + \delta_{k,s} - \delta_{k,s+1}$, we have

$$
\sum_k (j-k) + l_k(\mathcal{A}) = \begin{cases} 
\sum_k (j-k) + l_k(\mathcal{A}) + 1 & \text{if } r < j < s, \\
\sum_k (j-k) + l_k(\mathcal{A}) & \text{if } j \leq r - 1 \text{ or } j \geq s + 1.
\end{cases}
$$

Since all the Kronecker structure numbers other than $l_j$ of the pencils $\mathcal{A}$ and $\mathcal{C}$ coincide, it suffices to verify that

$$
(j-r)l_r(\mathcal{A}) = \sum_k (j-k) + l_k(\mathcal{A}) < \sum_k (j-k) + l_k(\mathcal{B}) \quad \text{if } r \leq j \leq s
$$

in order to show that $\mathcal{C} \prec \mathcal{B}$. When $r = s$, then the left-hand side of this inequality equals 0, while the right-hand side is not smaller then $(r-t)l_t(\mathcal{B}) > 0$; when $r < s$ then $l_r(\mathcal{A}) = 1$ and $(j-r)l_r(\mathcal{A}) = j - r < (j-t)l_t(\mathcal{B}) \leq \sum_{r}(j-k) + l_k(\mathcal{B})$, if only $r \leq j \leq s$. Thus in both subcases (12) holds true.

(ii) $r(\mathcal{A}) > 2$. Consider the dual pencils, which enter into case (i).

(iii) $l(\mathcal{A}) = 1$, $l(\mathcal{B}) > 2$. We represent the pencil $\mathcal{A}$ by a direct sum $\mathcal{A} = \mathcal{A}_0 \oplus \mathcal{L}_k$. Similarly to case (i), there exists an integer $i$ such that $0 < i < k$ and $l_i(\mathcal{B}) > 0$. We put $\mathcal{C} = \mathcal{A}_0 \oplus \mathcal{L}_{k-1} \oplus \mathcal{J}_i(\mu)$, where $\mu$ is not a Jordan eigenvalue of $\mathcal{A}$. The relation $\mathcal{A} \prec \mathcal{C}$ follows from Lemma 5, and the verification that $\mathcal{C} \prec \mathcal{B}$ is trivial.

(iv) $r(\mathcal{A}) = 1$, $r(\mathcal{B}) \geq 2$. One can pass to dual pencils and use (iii).

(v) $r(\mathcal{A}) = r(\mathcal{B}) = 1$, $l(\mathcal{A}) \leq 1$. We have, by (2), also $l(\mathcal{B}) = l(\mathcal{A}) \leq 1$. Since the proofs in the subcases $l(\mathcal{A}) = 1$, $l(\mathcal{A}) = 0$ are nearly the same, only the proof of the first case $l(\mathcal{A}) = l(\mathcal{B}) = 1$ is given. The pencils $\mathcal{A}$, $\mathcal{B}$ may be written as direct sums $\mathcal{A} = \mathcal{A}_0 \oplus \mathcal{L}_p \oplus \mathcal{L}_s$, $\mathcal{B} = \mathcal{B}_0 \oplus \mathcal{L}_r \oplus \mathcal{L}_t$. In view of (11), (8), and (9) we have $p > r$, $s > t$. The inequality (10) for $j$ large enough takes the form

$$
\sum_k k d_k(\mathcal{A}) \leq \sum_k k d_k(\mathcal{B})
$$

(13)

for all $\lambda \in \mathcal{C}$. We can assume that for some $\mu \in \mathcal{C}$ the inequality (13) is strict;
if such $\mu$ did not exist we would obtain a contradiction:

$$\dim X = p + s + 1 + \sum_{\lambda \in \mathcal{E}} \sum_{k} kd_k(\lambda, \mathcal{A})$$

$$\geq r + t + 1 + \sum_{\lambda \in \mathcal{E}} \sum_{k} kd_k(\lambda, \mathcal{B})$$

$$= \dim X.$$

Let $l$ be the smallest nonnegative integer such that $d_{l+j}(\mu, \mathcal{A}) = 0$ for $j = 1, 2, \ldots$. The pencil $\mathcal{A}$ may be identified with $\mathcal{A}_1 \oplus \mathcal{R}_p \oplus \mathcal{F}_l(\mu)$. One may easily verify that for the pencil $\mathcal{C} = \mathcal{A}_1 \oplus \mathcal{R}_{p-1} \oplus \mathcal{F}_{l+1}(\mu)$ we have $\mathcal{A} \preceq \mathcal{C} \preceq \mathcal{B}$.

(vi) $l(\mathcal{A}) = l(\mathcal{B}) = 1$, $r(\mathcal{A}) \leq 1$. One may use the argument of the previous case for the dual pencils.

Note that if the ordered pair of pencils $\mathcal{A} \prec \mathcal{B}$ does not fall into one of the above cases, then $\mathcal{A}$ is a regular pencil [i.e., $l(\mathcal{A}) = r(\mathcal{A}) = 0$] and $l(\mathcal{B}) = r(\mathcal{B})$. This is a consequence of (10) (for $\mu$ which is not a Jordan eigenvalue of $\mathcal{B}$ and $j = 1$) and (2).

(vii) $r(\mathcal{A}) = l(\mathcal{A}) = 0$, $l(\mathcal{B}) = r(\mathcal{B}) \geq 1$. Let $k(\lambda, \mathcal{B})$ denote the smallest nonnegative integer $k$ such that $d_{k+j}(\lambda, \mathcal{B}) = 0$ for all $j \geq 1$. We may assume that the pencil $\mathcal{B}$ has the following direct sum decomposition:

$$\mathcal{B} = \mathcal{B}_0 \oplus \bigoplus_{k = 1}^{r(\mathcal{B})} (\mathcal{L}_k \oplus \mathcal{B}_k).$$

Using (1) and (10), we obtain the inequality

$$\sum_{\lambda \in \mathcal{E}} \sum_{j > 0} jd_j(\lambda, \mathcal{B}) + \sum_{k = 1}^{r(\mathcal{B})} (i_k + j_k + 1) = \dim X = \sum_{\lambda \in \mathcal{E}} \sum_{j = 1}^{k(\lambda, \mathcal{A})} jd_j(\lambda, \mathcal{A})$$

$$\leq \sum_{\lambda \in \mathcal{E}} \sum_{j} jd_j(\lambda, \mathcal{B})$$

$$+ k(\lambda, \mathcal{A}) r(\mathcal{B}).$$

Hence

$$\sum_{k = 1}^{r(\mathcal{B})} (i_k + j_k + 1) \leq r(\mathcal{B}) \sum_{\lambda \in \mathcal{E}} k(\lambda, \mathcal{A}).$$
This inequality implies that for some $k$ we have $i_k + j_k + 1 \leq k(\lambda, \mathcal{A})$. Therefore there exist integers $p, q$ such that
\[ p \geq i_k, \quad q \geq j_k \tag{14} \]
and $p + q + 1 = \sum_{\lambda \in \mathcal{C}} k(\lambda, \mathcal{A})$. The pencil $\mathcal{A}$ may be written as $\mathcal{A} = \mathcal{A}_0 \oplus [\oplus_{\lambda \in \mathcal{C}} k(\lambda, \mathcal{A})(\lambda)]$. Setting $\mathcal{C} = \mathcal{A}_0 \oplus \mathcal{L}_p + \mathcal{R}_q$, we have $\mathcal{A} \prec \mathcal{C}$ by Lemma 5. Moreover, for any $\lambda \in \mathcal{C}$ and $j \leq k(\lambda, \mathcal{A})$ we have
\[ \sum_k \min(j, k) d_k(\lambda, \mathcal{A}) + fr(\mathcal{A}) - \sum_k \min(j, k) d_k(\lambda, \mathcal{A}) \leq \sum_k \min(j, k) d_k(\lambda, \mathcal{B}) + fr(\mathcal{B}), \]
while for $j > k(\lambda, \mathcal{A})$ we have
\[ \sum_k \min(j, k) d_k(\lambda, \mathcal{A}) + fr(\mathcal{A}) = \sum_k \min(j, k) d_k(\lambda, \mathcal{A}) + j - k(\lambda, \mathcal{A}) \leq \sum_k \min(k(\lambda, \mathcal{A}), k) d_k(\lambda, \mathcal{B}) + k(\lambda, \mathcal{A}) r(\mathcal{B}) + j - k(\lambda, \mathcal{A}) \leq \sum_k \min(j, k) d_k(\lambda, \mathcal{B}) + fr(\mathcal{B}). \]
These inequalities and (14) imply that $\mathcal{C} \prec \mathcal{B}$. 

(viii) $r(\mathcal{A}) = l(\mathcal{A}) = r(\mathcal{B}) = l(\mathcal{B}) = 0$. It follows from (10) that for each $\lambda \in \mathcal{C}$, $\sum_k kd_k(\lambda, \mathcal{A}) = \sum_k kd_k(\lambda, \mathcal{B})$. The sum of all the left-hand sides of these inequalities is equal to the dimension of $X$ and therefore is also equal to the sum of the right-hand sides over all $\lambda \in \mathcal{C}$. This implies that
\[ \sum_k kd_k(\lambda, \mathcal{A}) = \sum_k kd_k(\lambda, \mathcal{B}) \quad \text{for all} \quad \lambda \in \mathcal{C}. \tag{15} \]
We fix $\lambda \in \mathcal{C}$ such that this sum does not vanish. Writing (10) for $j = k(\lambda, \mathcal{A})$, we obtain the inequality
\[ \sum_k kd_k(\lambda, \mathcal{A}) \leq \sum_k \min(k, k(\lambda, \mathcal{A})) d_k(\lambda, \mathcal{B}). \]
This inequality with (15) and (11) implies that \( 1 \leq k(\lambda, \mathcal{S}) < k(\lambda, \mathcal{A}) \). Comparing this last inequality with (15), we conclude that \( \sum_{j=1}^{r} d_j(\lambda, \mathcal{S}) \geq 2 \).

Let \( r = k(\lambda, \mathcal{S}) \) and let \( p \) be the largest integer such that \( \sum_{p < j < r} d_j(\lambda, \mathcal{S}) \geq 2 \). We may assume that \( \mathcal{S} = \mathcal{S}_0 \otimes \mathcal{I}_p(\lambda) \otimes \mathcal{I}_r(\lambda) \); we put \( \mathcal{C} = \mathcal{S}_0 \otimes \mathcal{I}_{p-1}(\lambda) \otimes \mathcal{I}_{r+1}(\lambda) \). It follows from Lemma 5 that \( \mathcal{C} < \mathcal{S} \).

Note that [since \( d_j(\lambda, \mathcal{C}) = d_j(\lambda, \mathcal{S}) + \delta_{j,p-1} - \delta_{j,p} - \delta_{j,r} + \delta_{j,r+1} \)] we have

\[
\sum_{k} \min(k, j) [d_j(\lambda, \mathcal{C}) - d_j(\lambda, \mathcal{S})] = \begin{cases} -1 & \text{if } p < j \leq r, \\ 0 & \text{otherwise}. \end{cases}
\]

Thus in order to show that \( \mathcal{A} \preceq \mathcal{C} \) it suffices to verify that

\[
\sum_{k} \min(k, j) d_k(\lambda, \mathcal{A}) < \sum_{k} \min(k, j) d_k(\lambda, \mathcal{S}) \quad \text{when } p \leq j \leq r.
\]

If \( p = r \), then this inequality follows from (15) and \( r = k(\lambda, \mathcal{S}) < k(\lambda, \mathcal{A}) \). If \( p < r \) and \( p \leq j \leq r \), then \( d_j(\lambda, \mathcal{S}) = 1 \); using the previous arguments we may write

\[
\sum_{k} \min(k, j) d_k(\lambda, \mathcal{A}) \leq \sum_{k} kd_k(\lambda, \mathcal{A}) - [k(\lambda, \mathcal{A}) - j] d_k(\lambda, \mathcal{A}) < \sum_{k} kd_k(\lambda, \mathcal{S}) - (r - j) d_r(\lambda, \mathcal{S}) = \sum_{k} \min(j, k) d_k(\lambda, \mathcal{S}).
\]

This ends the proof.

**Proof of the "if" part of Theorem 1.** Given two ordered operator pencils \( \mathcal{A} \preceq \mathcal{B} \), putting \( \mathcal{A}_0 = \mathcal{A} \), \( \mathcal{B}_0 = \mathcal{B} \), we can construct, by Lemma 6, two sequences of pencils \( \{ \mathcal{A}_i \}, \{ \mathcal{B}_i \} \) such that \( \mathcal{A}_0 \preceq \mathcal{A}_{i+1} \preceq \mathcal{B}_{i+1} \preceq \mathcal{B}_i \) and either

\[
\mathcal{A}_i \preceq \mathcal{A}_{i+1} \quad \text{and} \quad \mathcal{B}_{i+1} = \mathcal{B}_i,
\]

or

\[
\mathcal{A}_i = \mathcal{A}_{i+1} \quad \text{and} \quad \mathcal{B}_{i+1} \preceq \mathcal{B}_i.
\]
It follows from Lemma 3 that these sequences are finite; therefore from elements of these sequences a sequence \( \{ C_i \}_{i=0}^n \) may be formed such that \( \mathcal{A} = C_0 < C_1 < \cdots < C_n = \mathcal{B} \). The transitivity of the relation \( \preceq \) implies that \( \mathcal{A} \preceq \mathcal{B} \).

Let us notice one application of Theorem 3. Given an operator \( A \) acting in a finite-dimensional space \( X \), we may consider a pencil \( \mathcal{A}(\lambda) = \mathcal{A} - \lambda I \). In this case the expression which appears in (10) has a simple interpretation:

\[
\sum_k \min\{j, d\} d_k(\lambda, \mathcal{A}) = \text{null}(A - \lambda I)^j.
\]

\( S(A) \), the similarity orbit of \( A \), is defined as the set of all the operators of the form \( S A S^{-1} \), where \( S \) is an invertible operator acting in \( X \). It follows from [3, Corollary 2.6 and Proposition 3.1] that an operator \( B \) belongs to the closure of \( S(A) \) if and only if \( \text{null}(B - \lambda I)^j \geq \text{null}(A - \lambda I)^j \) for all \( \lambda \in \mathbb{C} \) and all natural \( j \). Theorem 3 may be recognized as a generalization of this result.

MINIMAL PENCILS

An \((X, Y)\)-pencil \( \mathcal{M} \) is minimal in the sense of the relation \( \preceq \) if there exists no \((X, Y)\)-pencil \( \mathcal{C} \) such that \( \mathcal{C} \prec \mathcal{M} \). It follows from Lemma 3 that for any \((X, Y)\)-pencil \( \mathcal{A} \) there exists a minimal \((X, Y)\)-pencil \( \mathcal{M} \) such that \( \mathcal{M} \preceq \mathcal{A} \). The equivalence of the relations \( \preceq \) and \( \ll \) and the definition of \( \ll \) imply that the set of all minimal pencils is dense in the set of all \((X, Y)\)-pencils. The Kronecker canonical form of minimal pencils may be recognized with the help of Lemma 5. If \( \mathcal{M} \) is a minimal pencil, then claims (vi), (iii), and (iv) of that lemma imply that

\[
r(\mathcal{M}) l(\mathcal{M}) = r(\mathcal{M}) d_j(\lambda, \mathcal{M}) = l(\mathcal{M}) d_j(\lambda, \mathcal{M}) = 0 \quad (16)
\]

for all \( j \geq 1 \) and \( \lambda \in \overline{\mathbb{C}} \).

If \( \dim X > \dim Y \), then (2) and (16) imply that \( r(\mathcal{M}) = \dim X - \dim Y \), and claim (i) of Lemma 5 implies that \( \mathcal{M} \) is equivalent to \((I_k \otimes \mathcal{R}_j) \oplus (I_l \otimes \mathcal{R}_{j-1})\). The numbers \( j, k, l \) are easy to compute; using (1), we obtain the system of equations \((j + 1)k + jl = \dim X, k + l = r(\mathcal{M}) = \dim X - \dim Y \). A unique solution in nonnegative integers \( j, k, l \) \((l > 0) \) is \( j = E(\dim X / r(\mathcal{M})), k = \dim X - jr(\mathcal{M}), l = r(\mathcal{M}) - k \). It is worth noting that a minimal pencil
has the following simple matrix representation:

\[
\mathcal{M}(\lambda) = \begin{bmatrix}
-\lambda & 0 & \cdots & 0 & 1 \\
-\lambda & 0 & \cdots & 0 & 1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
-\lambda & 0 & \cdots & 0 & 1 \\
\end{bmatrix}
\]

If \(\dim X = \dim Y\), then (2) and (16) imply that \(r(\mathcal{M}) = l(\mathcal{M}) = 0\), and Lemma 5, claim (v), implies that the minimal pencil \(\mathcal{M}\) is equivalent to \(\bigoplus_{i=0}^{r} \mathcal{J}_{k_i}(\lambda_i)\), where \(\lambda_i \in \mathbb{C}\) are distinct points; and conversely any pencil of this form is minimal.

**Lemma 7.** The set of all minimal pencils is open.

**Sketch of the proof.** In the case when \(\dim X \neq \dim Y\) the thesis follows easily from Theorem 1.

Suppose that \(\dim X = \dim Y\) and that \(\mathcal{M}\) is a minimal \((X, Y)\)-pencil which is the limit of a sequence \(\mathcal{M}_n\) of nonminimal pencils. It follows from Theorem 1 that for \(n\) large the pencils \(\mathcal{M}_n\) are regular; passing to a subsequence we may assume moreover that \(\mathcal{M}_n\) are equivalent to \(\bigoplus_{i=0}^{r} \mathcal{J}_{k_i}(\mu_{i,n})\), \(k_i > 0\) \((0 \leq i \leq r)\), \(\mu_{0,n} = \mu_{1,n} \rightarrow \mu\), \(k_0 \leq k_1\). We define a pencil \(\mathcal{F}_n\) as in the proof of Theorem 2, in which \((\lambda_{1,n}, \lambda_{2,n}, \ldots, \lambda_{s,n}) = (\mu_{1,n}, \ldots, \mu_{1,n}, \mu_{2,n}, \ldots, \mu_{r,n})\) (each \(\mu_i\), except \(\mu_0\), is repeated \(k_i\) times), \(s = \dim X - k_0\). Then it is easy to verify that

\[
\text{null}(\mathcal{F}_n, \mathcal{A}_n) = \dim X,
\]

and \(\mathcal{F}_n\) converges to some \(\mathcal{F}\) for which \(\text{null}(\mathcal{F}, \mathcal{M}) = \dim X - k_0\). This leads to a contradiction, since for large \(n\) we should have

\[
\dim X = \text{null}(\mathcal{F}_n, \mathcal{A}_n) \leq \text{null}(\mathcal{F}, \mathcal{M}) = \dim X - k_0.
\]

In the case when \(\dim X = \dim Y > 1\) the set of all minimal pencils contains a proper subset which is also open and dense, namely the set of all regular pencils with simple \(J\)-eigenvalues (of multiplicity one) only, as is easily seen from the triangular form of a regular matrix pencil.

P. Van Dooren in [7] has observed that a computed Kronecker canonical form of a rectangular matrix pencil usually corresponds to the form of a pencil which we have called minimal. The fact that the set of all minimal pencils is open and dense in the set of all matrix pencils explains this observation.
REFERENCES


Received 31 July 1984; revised 7 October 1985