The Computation of Kronecker's Canonical Form of a Singular Pencil

P. Van Dooren
Division of Applied Mathematics and Computer Science
Katholieke Universiteit te Leuven
Louvain, Belgium

Submitted by Gene H. Golub

ABSTRACT

We develop stable algorithms for the computation of the Kronecker structure of an arbitrary pencil. This problem can be viewed as a generalization of the well-known eigenvalue problem of pencils of the type $\lambda I - A$. We first show that the elementary divisors $(\lambda - \alpha)^i$ of a regular pencil $\lambda B - A$ can be retrieved with a deflation algorithm acting on the expansion $(\lambda - \alpha)B - (A - \alpha B)$. This method is a straightforward generalization of Kublanovskaya's algorithm for the determination of the Jordan structure of a constant matrix. We also show how to use this method to determine the structure of the infinite elementary divisors of $\lambda B - A$. In the case of singular pencils, the occurrence of Kronecker indices—containing the singularity of the pencil—somewhat complicates the problem. Yet our algorithm retrieves these indices with no additional effort, when determining the elementary divisors of the pencil. The present ideas can also be used to separate from an arbitrary pencil a smaller regular pencil containing only the finite elementary divisors of the original one. This is shown to be an effective tool when used together with the QZ algorithm.

I. INTRODUCTION

The eigenvalue problem is one of the most discussed problems in linear algebra, both from the numerical point of view [1] and from the algebraic point of view [2].

It is well known that the problem of determining the eigenstructure (eigenvalues and associated Jordan structure) of an $n \times n$ matrix A is equivalent to finding the elementary divisors of the regular pencil $\lambda I_n - A$ [2]. Although the algebraic problem is well defined, the numerical problem can be very poorly conditioned (e.g. defective matrices). Several algorithms [3] have been elaborated for the eigenvalue problem, which are very satisfactory. Numerical (backward) stability is indeed guaranteed, even when the problem is ill posed.

LINEAR ALGEBRA AND ITS APPLICATIONS 27:103–140 (1979)

0024-3795/79/050103+38$01.75

However, when the eigenvalues of A are known, the problem of determining the Jordan structure is better posed than the general problem [4]. A computational algorithm directed towards this problem was first formulated by Kublanovskaya [5] and improved by Golub and Wilkinson [4].

The generalized eigenvalue problem for *regular* pencils $\lambda B - A$, though very similar to the previous problem, is more complicated, since infinite elementary divisors may also occur [2]. Numerical algorithms have been developed which deal efficiently with this problem [6, 7]. In this paper we show that when the eigenvalues of the pencil are known, a simple generalization of Kublanovskaya's algorithm yields analogous results with respect to stability and conditioning criteria. Also, this generalization is extended to treat *singular* pencils. For this case a canonical form was first presented by Kronecker [8]. However, current computational methods [2, 9, 10] are inefficient from the point of view of stability as well as that of algebraic complexity.

We show that the Kronecker indices, which are additional structure elements in the case of singular pencils, can be determined in much the same way as the infinite elementary divisors. This method can be used to separate in a stable way a regular pencil containing only the finite elementary divisors of the original singular pencil. If the finite eigenvalues (Smith zeros) are known, the algorithm can be continued to determine the complete structure. If they are not known, one may make use of the QZ algorithm on this smaller regular pencil.

During the past decade significant advances have been made in the theory of multivariable linear systems. In this area Kronecker's pencil theory has been applied to model matching [11], realization [12], linear feedback [13], modeling [14] and inversion [15, 16]. In each of these problems the structure of a singular pencil plays a significant role. Hence there is a need to determine the Kronecker structure in a reliable way. Several algorithms [2, 9, 10] are available for the determination of this structure, but they are numerically unstable. They indeed compute the Kronecker canonical form explicitly, but are then unstable, precluding pivoting techniques. As shown in this paper, one should in this case restrict oneself to the computation of a "staircase" form revealing the eigenstructure without putting the matrix in its canonical form.

During the review of this paper, Prof. Wilkinson drew our attention to an internal report [21] treating similar problems and also using similar techniques. The connection with the Kronecker structure, though, is not shown in his paper. We feel that it is exactly this feature that makes the picture more complete and that allows easier comprehension of the problem. The numerical approach and the treatment of finite eigenvalues are new.

In the following section we introduce our notation, and we review several results which we shall need in the sequel. In Sec. III the structure algorithm for regular pencils is explained, and we also show its relation to Kublanovskaya's algorithm. In Sec. IV we give the extension of the algorithm for singular pencils and comment on its use for the extraction of the "regular part" of a singular pencil. In Sec. V we mention several applications, such as in the area of linear systems.

II. PRELIMINARIES

Throughout the paper we use uppercase for matrices (I for the unit matrix) and lowercase for vectors and scalars. Greek letters are used in conformity with standard conventions in the literature. All matrices, vectors and functions considered are defined over $\mathbb{C}$. By A^* we denote the conjugate transpose of A, and by A^P the "pertranspose" of A, which is the transpose over the second diagonal.

In our exposition we frequently use invertible row transformations to reduce an arbitrary $m\times n$ matrix A to the form

$$U^*\cdot A=\left[\begin{array}{c} A_r \\ \hline 0 \end{array}\right]\}\rho, \tag{2.1}$$

wherein A_r has ρ linearly independent rows (ρ is then clearly the rank of A). We call such a transformation a "row compression" of the matrix A. Analogously we use the name "column compression" for the invertible column transformation

$$A\cdot V=[\underbrace{A_c}_{\rho}\mid 0], \tag{2.2}$$

wherein the columns of A_c are linearly independent. The resulting matrices A_r and A_c are said to have zero row nullity (or full row rank) and zero column nullity (or full column rank), respectively. Therefore A_r has a right inverse A_r^+ and A_c a left inverse A_c^+ satisfying

$$A_rA_r^+=I_\rho=A_c^+A_c \tag{2.3}$$

We reserve the specification "full rank" for square invertible matrices. A reliable way of computing these compressions is by making use of the singular-value decomposition [3, 4] of the $m\times n$ matrix A:

$$A=U\cdot\Sigma\cdot V^*,$$

where

(1) U and V are respectively $m\times m$ and $n\times n$ unitary matrices,
(2) Σ is a $m\times n$ matrix of the form

$$\Sigma=\left[\begin{array}{c|c}\Sigma_\rho & 0\\ \hline 0 & 0\end{array}\right],\qquad \Sigma_\rho=\operatorname{diag}\{\sigma_1,\ldots,\sigma_\rho\},$$

with σ_i being positive real and satisfying $\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_\rho>0$.

It may be readily verified that U^*A and AV yield respectively a row compression and a column compression of A.

The matrix function $\lambda B-A$, with A and B arbitrary constant matrices of equal dimensions, is called a regular pencil when $\det(\lambda B-A)\not\equiv 0$ and a singular pencil otherwise. Regular pencils are thus always square.

DEFINITION. Two pencils λB_1-A_1 and λB_2-A_2 of dimension $m\times n$ are said to be *strictly equivalent* when there exist constant invertible matrices P and Q of orders m and n respectively, such that

$$P(\lambda B_1-A_1)Q=\lambda B_2-A_2 \tag{2.4}$$

We will denote this equivalence relation by $\sim$.

Kronecker's theory of singular pencils [2] shows that any pencil $\lambda B-A$ has, under strict equivalence, a canonical quasidiagonal form:

$$P(\lambda B-A)Q=\operatorname{diag}\{L_{\varepsilon_1},\ldots,L_{\varepsilon_p},L^P_{\eta_1},\ldots,L^P_{\eta_q},\lambda N-I,\lambda I-J\}, \tag{2.5}$$

where

(1) L_μ is the $\mu\times(\mu+1)$ bidiagonal pencil

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

(2) L^P_μ is the "pertransposed" pencil

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

of dimensions $(\mu+1)\times\mu$,

(3) N is a nilpotent Jordan matrix, and
(4) J is in Jordan canonical form.

Hence, $\lambda I - J$ contains the finite elementary divisors and $\lambda N - I$ the infinite elementary divisors. Also, the blocks L_{ε_i} and $L_{\eta_j}^{P}$ contain the singularity of the pencil, since there exist polynomial (column or row) vectors that zero out these blocks identically:

$$\mu\left\{\underbrace{\begin{bmatrix} \lambda & -1 & & \\ & \ddots & \ddots & \\ & & \lambda & -1 \end{bmatrix}}_{\mu+1}\right.\begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \\ \vdots \\ \lambda^{\mu} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{2.6}$$

The sizes of these blocks characterize them completely and are therefore given specific names. The ε_i are called Kronecker column indices and the η_j are called Kronecker row indices.

The pencil $\lambda N - I$ is completely determined by the degrees δ_i of the infinite elementary divisors, and $\lambda I - J$ by the finite elementary divisors $(\lambda - \alpha_i)^{j}$.

Example 1. A possible canonical form is e.g. the 14 by 16 pencil.

$$\left[\begin{array}{cccccccc}
\quad & \begin{array}{|cc|}\hline \lambda & -1 \\ \hline\end{array} & & & & & & \\
& & \begin{array}{|ccc|}\hline \lambda & -1 & \\ & \lambda & -1 \\ \hline\end{array} & & & & & \\
& & & \begin{array}{|ccc|}\hline -1 & & \\ \lambda & -1 & \\ & \lambda & -1 \\ & & \lambda \\ \hline\end{array} & & & & \\
& & & & \begin{array}{|c|}\hline -1 \\ \hline\end{array} & & & \\
& & & & & \begin{array}{|cc|}\hline -1 & \\ \lambda & -1 \\ \hline\end{array} & & \\
& & & & & & \begin{array}{|c|}\hline \lambda-2 \\ \hline\end{array} & \\
& & & & & & & \begin{array}{|cc|}\hline \lambda-3 & \\ -1 & \lambda-3 \\ \hline\end{array}
\end{array}\right] \tag{2.7}$$

Then

(i) the Kronecker indices are

$$\varepsilon_1=0,\quad \varepsilon_2=0,\quad \varepsilon_3=1,\quad \varepsilon_4=2,$$

$$\eta_1=0,\quad \eta_2=3;$$

(ii) there are two infinite elementary divisors: $\delta_1=1$, $\delta_2=2$;
(iii) the finite elementary divisors are $\lambda-2$ and $(\lambda-3)^2$.

The elements summed up in (i), (ii) and (iii) completely characterize the canonical form (2.7).

As shown in (2.6), one can always find a polynomial vector y of degree μ to zero out the block L_μ. By appropriately completing y with zeros, one finds a corresponding "right null vector" for the canonical pencil containing such a L_μ block. Similar remarks apply to left null vectors; in the example (2.7) these are

$$\begin{array}{ll} [0\ \ 0\ \ 0\ \ 1\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0] & \text{corresponding to } L^P_{\eta_1}, \\ [0\ \ 0\ \ 0\ \ 0\ \ \lambda^3\ \ \lambda^2\ \ \lambda\ \ 1\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0] & \text{corresponding to } L^P_{\eta_2}. \end{array}$$

Zero Kronecker indices thus correspond to constant null vectors (degree $\mu=0$).

A Kronecker block clearly has deficient column (or row) rank for *any* value of λ, and a corresponding constant null vector can be found by substituting that value of λ in the corresponding equation (2.6).

Regular pencils $\lambda B-A$ have no Kronecker indices, since $\det(\lambda B-A)\not\equiv 0$ and not all the values of λ will yield rank deficiency of the matrix. We prove in Sec. IV that it is always possible, by a deflation procedure using unitary transformations only, to separate from a *singular* pencil $\lambda B-A$ a regular pencil λB_r-A_r with the same elementary divisors as $\lambda B-A$. The canonical form of λB_r-A_r is thus

$$\lambda B_r-A_r \sim \operatorname{diag}\{\lambda N-I, \lambda I-J\} \tag{2.8}$$

with N and J as in (2.5).

If the regular pencils $\lambda I-A_1$ and $\lambda I-A_2$ are strictly equivalent, then A_1 and A_2 are similar (since $P=Q^{-1}$). The Kronecker canonical form can therefore be viewed as a generalization of the Jordan canonical form to arbitrary pencils. At a first level of generalization are the regular pencils $\lambda B-A$ (bringing in infinite elementary divisors), and at a second level are the arbitrary pencils (introducing the Kronecker indices).

While the link with the eigenvalue problem is obvious for pencils $\lambda I - A$, the concept of eigenvalue is rather ambiguous for arbitrary pencils (different definitions occur in the literature). In analogy to infinite and finite elementary divisors, we connect the terms "infinite" and "finite" eigenvalue with regular pencils and with the "regular part" $\lambda B_r - A_r$ of singular pencils as well. We thus say that the pencil (2.7) has three infinite eigenvalues: two eigenvalues at 3 and one at 2.

We use the name "finite structure" for the set of finite elementary divisors (or Jordan information of $\lambda I - J$), and "infinite structure" for the set of infinite elementary divisors (or Jordan information of $\lambda N - I$). The set of Kronecker indices will be called the "singular structure" of the pencil. These three structure elements thus completely determine the canonical form of a pencil and conversely.

An algorithm for the computation of the Jordan information of a pencil of the type $\lambda I - A$ was developed by Kublanovskaya, in the case where the eigenvalues of A are known. It operates on the constant term of the expansion $(\lambda - \alpha)I - (A - \alpha I)$. Clearly $\lambda I - A$ has elementary divisors at α when $A - \alpha I$ is singular. Moreover, the nullities ν_j of $(A - \alpha I)^j$ for $j = 1, \ldots, l$ contain the Jordan information of A at α [2] (where ν_j does not change any more after l). The following algorithm computes these nullities in a stable way [4, 5] (comments are given between *comment* and *;*).

Algorithm 2.1.

```
        comment initialization ;
        $j := 1$; $A_{1,1} := A - \alpha I$; $n_1 := n$;
step _ j : comment compute the singular value decomposition of the $n_j \times n_j$
        matrix $A_{j,j}$ ;
        call SVD $(A_{j,j})$ result $(U_j, \Sigma_j, V_j)$ rank $(r_j)$ nullity $(s_j)$;
        if $s_j = 0$ then begin $l := j - 1$; stop end;
        comment compress $A_{j,j}$ to full column rank $r_j$ and partition ;

        $\left[\begin{array}{c|c} A_{j+1,j+1} & 0 \\ \hline A_{j+1,j} & 0 \end{array}\right] := V_j^* A_{j,j} V_j$;

        comment also update and partition blocks in column $j$ ;
        for $i = 1$ step 1 until $j - 1$ do $\left[\, A_{j+1,i} \mid A_{j,i} \,\right] := A_{j,i} V_j$;
        comment update ; $n_{j+1} := r_j$; $j := j + 1$; go to step_j ;
```

The nullities ν_j, $j = 1, \ldots, l$, are then given by [4]:

$$\nu_j = \sum_{i=1}^{j} s_i. \tag{2.9}$$

This algorithm implicitly computes a similarity transformation to put $\lambda I - A$ in a "staircase" form. When one embeds each V_j in a $n \times n$ matrix and puts

$$V = \prod_{j=1}^{l} \left[\begin{array}{c|c} V_j & \\ \hline & I \end{array}\right],$$

the matrix $\hat{A} = V^*(A - \alpha I)V$ has the form (for some X)

$$\hat{A} = \left[\begin{array}{c|c} A_{l+1,l+1} & \\ \hline X & A_\alpha \end{array}\right] \begin{array}{l} \}r_l \\ \}\Delta = \sum_{j=1}^{l} s_j \end{array}$$

where $A_{l+1,l+1}$ has full rank and A_α is zero on and above the main diagonal (A_α is thus nilpotent). This yields the structure of A at α, since V^*AV has the form

$$V^*AV = \hat{A} + \alpha I = \left[\begin{array}{c|c} A_{l+1,l+1} + \alpha I_{r_l} & 0 \\ \hline X & A_\alpha + \alpha I_\Delta \end{array}\right] = \left[\begin{array}{c|c} \tilde{A} & 0 \\ \hline X & A_\alpha + \alpha I_\Delta \end{array}\right], \tag{2.10}$$

where

$$A_\alpha + \alpha I_\Delta = \left[\begin{array}{c|c|c|c|c} \alpha I_{s_l} & & & & \\ \hline \vdots & \ddots & & & \\ \hline A_{l,3} & \cdots & \alpha I_{s_3} & & \\ \hline A_{l,2} & \cdots & A_{3,2} & \alpha I_{s_2} & \\ \hline A_{l,1} & \cdots & A_{3,1} & A_{2,1} & \alpha I_{s_1} \end{array}\right]. \tag{2.11}$$

As a consequence of the rank search in Algorithm 2.1, the off-diagonal blocks $A_{i+1,i}$ have full column rank s_i. Therefore, the s_i are decreasing with i (remark that $s_{l+1} = 0$ because of the stopping rule of the algorithm). This implies [4] that $A_\alpha + \alpha I_\Delta$ and A have the following Jordan structure:

$$A \text{ has } s_j - s_{j+1} = a_j \text{ Jordan chains of size } j \qquad (\text{for } \quad j = l, \dots, 1). \tag{2.12}$$

As expected from (2.9), the index set $\{s_i\}$ thus also yields the Jordan structure of A at α.

Since $A_{l+1,l+1}$ in (2.10) has full rank, $A_{l+1,l+1}+\alpha I_{\tilde{r}_l}=\tilde{A}$ has no eigenvalue at α. The procedure is then restarted on $\lambda I_{\tilde{r}_l}-\tilde{A}$ using an expansion at another eigenvalue. This process is continued until the total structure of A is recovered. Each step of this algorithm can also be viewed as a deflation step whereby s_l eigenvalues α are retrieved at a time. When no eigenvalues at α are left, the algorithm is restarted at another point.

In the following section we present an analogous rank search for regular pencils of the more general type $\lambda B-A$.

III. REGULAR PENCILS

When computing the eigenstructure of the regular pencil $\lambda I-A$, one will use similarity transformations in order to preserve the I matrix of the original pencil (see Algorithm 2.1). On the other hand, when dealing with the more general form $\lambda B-A$, one has the possibility of making use of independent left and right multiplications. In this section we show how this additional degree of freedom enables us to transform both A and B into specific "staircase" forms, yielding the required structure information. Before describing the actual algorithm, we work out a few steps of it in detail in order to clarify the procedure.

The aim is to retrieve the structure of the eigenvalue α of the *regular* pencil $\lambda B-A$. Since α is an eigenvalue, the constant term of the expansion at α,

$$(\lambda-\alpha)B-(A-\alpha B), \tag{3.1}$$

is singular. For simplicity rename $\lambda-\alpha$, B and $A-\alpha B$ respectively as λ', $B_{1,1}$ and $A_{1,1}$. We thus have an expansion

$$\lambda' B_{1,1}-A_{1,1} \tag{3.2}$$

with $A_{1,1}$ singular. The transformation $\lambda-\alpha=\lambda'$ is also called a "shift" whereby the eigenvalue α of (3.1) is transformed (or "shifted") to an eigenvalue 0 of the pencil (3.2). Using this shift, all finite eigenvalues can thus be treated similarly. We then perform following steps:

Step 1

(a) Compute a singular-value decomposition of $A_{1,1}$,

$$A_{1,1}=U_a\Sigma_a V_a^* \qquad \text{(with nullity } s_1\text{)},$$

and multiply the pencil by V_a on the right. Hence the columns of $A_{1,1}$ are "compressed" in the first $n-s_1$ columns, and the remaining s_1 columns are zero. The resulting pencil is then partitioned as

$$(\lambda' B_{1,1}-A_{1,1})V_a=[\underbrace{\lambda' B_2-A_2}_{n-s_1} \mid \underbrace{\lambda' B_1}_{s_1}]\}n. \tag{3.3}$$

(b) Compute a singular-value decomposition of B_1,

$$B_1=U_b\Sigma_b V_b^*,$$

and multiply the pencil left by U_b^* on the left, which compresses B_1 to full row rank. The rank must be s_1, since if it were less, that would imply that $\det(\lambda B-A)=0$; hence we can partition as follows:

$$U_b^*(\lambda' B_{1,1}-A_{1,1})V_a=\left[\begin{array}{c|c} \lambda' B_{2,1}-A_{2,1} & \lambda' B_{1,1} \\ \hline \lambda' B_{2,2}-A_{2,2} & 0 \end{array}\right]\begin{array}{l}\}s_1 \\ \}n-s_1\end{array}. \tag{3.4}$$
$$\qquad\qquad\qquad\qquad\underbrace{\qquad\quad}_{n-s_1}\quad\underbrace{\quad}_{s_1}$$

(c) Permute the top and bottom blocks (this is a unitary row operation P_b). Hence we have constructed *unitary* transformations P_1 and Q_1 such that

$$P_1(\lambda' B_{1,1}-A_{1,1})Q_1=\left[\begin{array}{c|c} \lambda' B_{2,2}-A_{2,2} & 0 \\ \hline \lambda' B_{2,1}-A_{2,1} & \lambda' B_{1,1} \end{array}\right]\begin{array}{l}\}n_2 \\ \}s_1\end{array}, \tag{3.5}$$
$$\qquad\qquad\qquad\qquad\underbrace{\qquad\quad}_{n_2}\quad\underbrace{\quad}_{s_1}$$

where

(i) $B_{1,1}$ has full rank s_1 (this is the first separated "stair"),

(ii) $\left[\begin{array}{c} A_{2,2} \\ \hline A_{2,1} \end{array}\right]$ has full column rank n_2.

Note that we reuse some of the names of the blocks—e.g. $B_{1,1}$—after processing them. This is done to avoid step indices. Just as in the Kublanovskaya algorithm, this step may be viewed as a deflation. The eigenvalue problem (3.5) indeed splits up into two eigenvalue problems $\lambda' B_{1,1}$ and $\lambda' B_{2,2}-A_{2,2}$. The first pencil clearly has all its eigenvalues at $\lambda'=0$, since $\det(\lambda' B_{1,1})=(\lambda')^{s_1}\det(B_{1,1})$. We have thus reduced the problem to the processing of $\lambda' B_{2,2}-A_{2,2}$. While the off-diagonal blocks $\lambda' B_{2,1}-A_{2,1}$ do not play a role in the eigenvalue problem, they do when the Jordan structure of the eigenvalues are required. The off-diagonal blocks will therefore also be processed in the algorithm.

Step 2:

(a), (b), (c) Repeat the above procedure on $\lambda' B_{2,2}-A_{2,2}$ with $n_2 \times n_2$ matrices $\hat{P}_2$ and $\hat{Q}_2$, thus obtaining a similar reduction of this smaller pencil. These transformations may be embedded in

$$P_2=\left[\begin{array}{c|c} \hat{P}_2 & \\ \hline & I_{s_1} \end{array}\right] \quad \text{and} \quad Q_2=\left[\begin{array}{c|c} \hat{Q}_2 & \\ \hline & I_{s_1} \end{array}\right], \tag{3.6}$$

so that we have

$$P_2P_1(\lambda' B_{1,1}-A_{1,1})Q_1Q_2 = \begin{array}{cc} \left[\begin{array}{c|c} \hat{P}_2(\lambda' B_{2,2}-A_{2,2})\hat{Q}_2 & 0 \\ \hline (\lambda' B_{2,1}-A_{2,1})\hat{Q}_2 & \lambda' B_{1,1} \end{array}\right] & \begin{array}{l} \}n_2 \\ \}s_1 \end{array} \\ \begin{array}{cc} \underbrace{\hphantom{\hat{P}_2(\lambda' B_{2,2}-A_{2,2})\hat{Q}_2}}_{n_2} & \underbrace{\hphantom{\lambda' B_{1,1}}}_{s_1} \end{array} & \end{array}$$

$$= \begin{array}{cc} \left[\begin{array}{c|c|c} \lambda' B_{3,3}-A_{3,3} & 0 & 0 \\ \hline \lambda' B_{3,2}-A_{3,2} & \lambda' B_{2,2} & 0 \\ \hline \lambda' B_{3,1}-A_{3,1} & \lambda' B_{2,1}-A_{2,1} & \lambda' B_{1,1} \end{array}\right] & \begin{array}{l} \}n_3 \\ \}s_2 \\ \}s_1 \end{array}, \\ \begin{array}{ccc} \underbrace{\hphantom{\lambda' B_{3,3}-A_{3,3}}}_{n_3} & \underbrace{\hphantom{\lambda' B_{2,1}-A_{2,1}}}_{s_2} & \underbrace{\hphantom{\lambda' B_{1,1}}}_{s_1} \end{array} & \end{array} \tag{3.7}$$

where

(i) $B_{2,2}$ and $B_{1,1}$ have full rank,

(ii) $A_{2,1}$ and $\left[\begin{array}{c} A_{3,3} \\ \hline A_{3,2} \end{array}\right]$ have full column rank.

(Again we reuse some of the names of the blocks.)

Step j (induction step)

Repeat this procedure until the upper left block $A_{j,j}$ has full rank.

This procedure is recapitulated in the following in ALGOL-type notation (comments are given between *comment* and ;):

```
ALGORITHM 3.1.
    comment initialization ;
    $j := 1;\ A_{1,1} := A - \alpha B;\ B_{1,1} := B;\ n_1 := n;$
step _ $j$: comment compute the S.V.D. of the $n_j \times n_j$ matrix $A_{j,j}$ ;
    call SVD $(A_{j,j})$ result $(U_a, \Sigma_a, V_a)$ nullity $(s_j)$;
    if $s_j = 0$ then begin $l := j - 1$; stop end;
    comment compress $A_{j,j}$ to full column rank and partition ;
    $\left[\begin{array}{c|c} A_{j+1} & 0 \end{array}\right] := A_{j,j}V_a;\ \left[\begin{array}{c|c} B_{j+1} & B_j \end{array}\right] := B_{j,j}V_a;$
    comment also update and partition blocks in column $j$ ;
    for $i = 1$ step 1 until $j - 1$ do
        begin $\left[\begin{array}{c|c} A_{j+1,i} & A_{j,i} \end{array}\right] := A_{j,i}V_a;\ \left[\begin{array}{c|c} B_{j+1,i} & B_{j,i} \end{array}\right] := B_{j,i}V_a$ end;
    comment compute the S.V.D. of the $n_j \times s_j$ matrix $B_j$ ;
    call SVD $(B_j)$ result $(U_b, \Sigma_b, V_b)$ rank $(s_j)$;
    comment compress $B_j$ to full row rank, permute and partition ;
    $\left[\begin{array}{c} A_{j+1,j+1} \\ \hline A_{j\ +1,j} \end{array}\right] := P_b U_b^* A_{j+1};\ \left[\begin{array}{c} B_{j+1,j+1} \\ \hline B_{j+1,j} \end{array}\right] := P_b U_b^* B_{j+1};\ \left[\begin{array}{c} 0 \\ \hline B_{j,j} \end{array}\right] := P_b U_b^* B_j;$
    comment update ; $n_{j+1} := n_j - s_j;\ j := j + 1;$ go to step_$j$;
```

Note that the stopping rule of the algorithm is when $A_{j,j}$ has full column rank (then $s_j = 0$ and $l = j - 1$). This algorithm reduces $\lambda' B_{1,1} - A_{1,1}$ to the form (for some X)

$$P(\lambda B - A)Q = P(\lambda' B_{1,1} - A_{1,1})Q = \left[\begin{array}{c|c} \lambda' B_{l+1,l+1} - A_{l+1,l+1} & 0 \\ \hline X & \lambda' B_\alpha - A_\alpha \end{array}\right]$$

$$= \begin{array}{c} \left[\begin{array}{c|c|c|c|c} \lambda' B_{l+1,l+1} - A_{l+1,l+1} & 0 & \cdots & 0 & 0 \\ \hline \lambda' B_{l+1,l} - A_{l+1,l} & \lambda' B_{l,l} & \cdots & 0 & 0 \\ \hline \vdots & \vdots & \ddots & \vdots & \vdots \\ \hline \lambda' B_{l+1,2} - A_{l+1,2} & \lambda' B_{l,2} - A_{l,2} & \cdots & \lambda' B_{2,2} & 0 \\ \hline \lambda' B_{l+1,1} - A_{l+1,1} & \lambda' B_{l,1} - A_{l,1} & \cdots & \lambda' B_{2,1} - A_{2,1} & \lambda' B_{1,1} \end{array}\right] \begin{array}{l} \}n_{l+1} \\ \}s_l \\ \\ \}s_2 \\ \}s_1 \end{array} \\ \begin{array}{ccccc} \underbrace{\hspace{3cm}}_{n_{l+1}} & \underbrace{\hspace{2cm}}_{s_l} & \qquad & \underbrace{\hspace{2cm}}_{s_2} & \underbrace{\hspace{1cm}}_{s_1} \end{array} \end{array}, \tag{3.8}$$

where

(1) $A_{l+1,l+1}$ has full rank (thus $s_{l+1} = 0$),
(2) the $B_{i,i}$ have full rank s_i for $i = 1, \ldots, l$,
(3) the $A_{i,i-1}$ have full column rank s_i for $i = 2, \ldots, l$.

From this last property it follows that the s_i form a decreasing sequence and thus

$$s_j - s_{j+1} = a_j \geqslant 0 \qquad \text{for} \quad j = l, \ldots, 1. \tag{3.9}$$

The analogy between Algorithms 3.1 and 2.1 suggests that these indices again contain information about the structure of $\lambda B - A$ at α. In Lemmas 3.2, 3.3 and 3.4 and Proposition 3.5 we prove that the form (3.8) indeed yields the Jordan structure of $\lambda B - A$ at α. Each step of the reasoning brings the form (3.8) closer to its canonical form. Since the transformations used for this purpose are not unitary, proceeding beyond the form (3.8) is not recommended.

LEMMA 3.2. *The pencil* (3.8) *is strictly equivalent to the pencil*

$$\operatorname{diag}\{\lambda' B_{l+1,l+1} - A_{l+1,l+1}, \lambda' B_\alpha - A_\alpha\}. \tag{3.10}$$

Proof. We prove this inductively by constructing the transformations that zero the blocks $A_{l+1,i}$ and $B_{l+1,i}$ for $i = l, \ldots, 1$. The blocks used thereby as pivots are left or right invertible according to (2.3).

Step l. Since $A_{l+1,l+1}$ has full (column) rank, there exists a row transformation that uses this block to zero out $A_{l+1,l}$ (see Fig. 1). Afterwards the modified $B_{l+1,l}$ is eliminated by a column transformation using $B_{l,l}$ as pivot, since this block has full (row) rank.

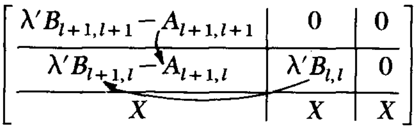

FIG. 1.

Induction step i. Having already eliminated $\lambda' B_{l+1,j} - A_{l+1,j}$ for $j > i$, we perform the elimination of $\lambda' B_{l+1,i} - A_{l+1,i}$ in an analogous fashion. This is done by using $A_{l+1,l+1}$ and $B_{i,i}$ as pivots, as in Fig. 2.

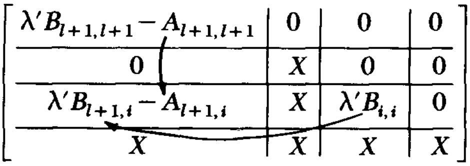

FIG. 2

Note that none of these transformations affect the diagonal blocks of (3.10). ∎

LEMMA 3.3. *The pencil* $\lambda' B_\alpha - A_\alpha$ *in* (3.10) *is strictly equivalent to the bidiagonal block pencil*

$$\lambda' B_{bi} - A_{bi} = \begin{bmatrix} \lambda' B_{l,l} & & & 0 \\ -A_{l,l-1} & \ddots & & \\ & \ddots & \ddots & \\ 0 & & -A_{2,1} & \lambda' B_{1,1} \end{bmatrix}. \tag{3.11}$$

Proof. We show by induction that in each column j, all elements other than the blocks $\lambda' B_{j,j}$ and $A_{j,j-1}$ can be eliminated. Suppose that columns $l,\ldots,j+1$ have already been treated. We first eliminate $B_{j,j-1}$ using $B_{j-1,j-1}$ as pivot. Here $A_{j,j-1}$ remains unchanged. The situation is then as shown in Fig. 3.

$$\left[\begin{array}{ccc|c:ccc}
X & & & & & & \\
X & \ddots & & & & & \\
 & \ddots & X & & & & \\
 & & X & \lambda' B_{j,j} & & & \\ \hline
 & & & -A_{j,j-1} & \lambda' B_{j-1,j-1} & & \\
 & & & & & \ddots & \\
 & & & X & X & & \lambda' B_{1,1}
\end{array}\right]$$

FIG. 3

Since $A_{j,j-1}$ has full column rank, we have in the bottom right corner a situation similar to that treated in Lemma 3.2. Hence we may use the method of this previous lemma to zero out all information under $A_{j,j-1}$. Observe that these transformations do not affect the blocks $B_{i,i}$ and $A_{i,i-1}$. ∎

Denote

$$J_i = I_{s_i} \quad \text{and} \quad K_i = \underbrace{\left[\begin{array}{c} I_{s_{i+1}} \\ \hline 0 \end{array}\right]}_{s_{i+1}} \Big\} s_i. \tag{3.12}$$

Lemma 3.4. *The bidiagonal pencil* $\lambda' B_{bi} - A_{bi}$ *in* (3.11) *is strictly equivalent to the normalized pencil*

$$\lambda' B_n - A_n = \begin{bmatrix} \lambda' J_l & & & & 0 \\ -K_{l-1} & \cdot & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ 0 & & & -K_1 & \lambda' J_1 \end{bmatrix}. \tag{3.13}$$

Proof. Since in $\lambda' B_{bi} - A_{bi}$ the matrix $B_{l,l}$ has full (row) rank s_l, there exists a column transformation that reduces it to the form J_l. This operation preserves the full column rank, s_l, of $A_{l,l-1}$. Also there exists a row transformation "normalizing" $A_{l,l-1}$ to K_{l-1}, though affecting $B_{l-1,l-1}$. This process of normalization can be continued throughout the whole matrix and yields the desired result. ■

These three lemmas and our algorithm thus construct the equivalences

$$\lambda' B_{1,1} - A_{1,1}$$

$$\sim \left[\begin{array}{c|c} \lambda' B_{l+1,l+1} - A_{l+1,l+1} & 0 \\ \hline X & \lambda' B_\alpha - A_\alpha \end{array}\right] \sim \left[\begin{array}{c|c} \lambda' B_{l+1,l+1} - A_{l+1,l+1} & 0 \\ \hline 0 & \lambda' B_\alpha - A_\alpha \end{array}\right] \tag{3.14a}$$

and

$$\lambda' B_\alpha - A_\alpha \sim \lambda' B_{bi} - A_{bi} \sim \lambda' B_n - A_n. \tag{3.14b}$$

As a consequence we have

Proposition 3.5. *The indices* $\{s_i\}$ *given by Algorithm* 3.1 *completely determine the structure at* α *of the pencil* $\lambda B - A$:

$$\lambda B - A \text{ has } (s_j - s_{j+1}) = a_j \text{ elementary divisors } (\lambda - \alpha)^j \qquad (j = 1, \ldots, l).$$

Proof. From (3.14) it is easy to see that $\lambda' B_{1,1} - A_{1,1}$ is strictly equivalent to $\operatorname{diag}\{\lambda' B_{l+1,l+1} - A_{l+1,l+1}, \lambda' B_n - A_n\}$. Keeping in mind the transformation $\lambda-\alpha=\lambda'$, this also gives

$$(\lambda-\alpha)B-(A-\alpha B)$$

$$\sim\left[\begin{array}{c|cccc}
(\lambda-\alpha)B_{l+1,l+1}-A_{l+1,l+1} & & & & \\
\hline
 & (\lambda-\alpha)J_l & & & \\
 & -K_{l-1} & \ddots & & \\
0 & & \ddots & \ddots & \\
 & & & -K_1 & (\lambda-\alpha)J_1
\end{array}\right]$$

$$(3.15)$$

Since $A_{l+1,l+1}$ has full rank, the top pencil has no eigenvalues at α. The bottom pencil (separated by Algorithm 3.1) has all its eigenvalues at α, and its structure is easy to retrieve. This is indeed a pencil of the type $\lambda I-A$ with A in quasi-Jordan form. According to (2.12) this pencil has

$$(s_j-s_{j+1})=a_j \text{ Jordan chains of size } j \qquad (\text{for} \quad j=l,\dots,1) \qquad ■$$

REMARK. It is easy to see that Algorithm 3.1 is essentially Kublanovkaya's algorithm when applied to the regular $n\times n$ pencil $\lambda I-A$. Since the B matrix is the identity matrix in this situation, the second compression (done by $P_b U_b^*$ in Algorithm 3.1) is done by V^* in Algorithm 2.1. Because of this choice of similarity transformations, B remains the identity matrix (which also appreciably reduces the amount of work).

Hence the reduced matrix is similar to (2.10):

$$(\lambda-\alpha)I-(A-\alpha I)$$

$$\left[\begin{array}{c|c|c|c|c}
(\lambda-\alpha)I_{n_{l+1}}-A_{l+1,l+1} & 0 & \cdots & 0 & 0 \\
\hline
X & (\lambda-\alpha)I_{s_l} & \cdots & 0 & 0 \\
\hline
\vdots & \vdots & & \vdots & \vdots \\
\hline
X & X & \cdots & (\lambda-\alpha)I_{s_2} & 0 \\
\hline
X & X & \cdots & -A_{2,1} & (\lambda-\alpha)I_{s_1}
\end{array}\right].$$

$$■ \quad (3.16)$$

Algorithm 3.1 applied to the eigenvalue α of $\lambda B-A$ thus "deflates" the structure of the pencil at α. In the next step one proceeds on the deflated pencil $(\lambda-\alpha)B_{l+1,l+1}-A_{l+1,l+1}$ with a second eigenvalue. After processing all the finite eigenvalues we end up with

$$P(\lambda B-A)Q=\left[\begin{array}{c|c|c|c} \lambda\tilde{B}-\tilde{A} & 0 & \cdots & 0 \\ \hline X & (\lambda-\alpha_k)B_{\alpha_k}-A_{\alpha_k} & \cdots & 0 \\ \hline \vdots & \vdots & & \vdots \\ \hline X & X & \cdots & (\lambda-\alpha_1)B_{\alpha_1}-A_{\alpha_1} \end{array}\right], \tag{3.17}$$

where $(\lambda-\alpha_i)B_{\alpha_i}-A_{\alpha_i}$ has a specific "staircase" structure [see (3.8)] yielding indices $\{s_j\}$ for the eigenvalue α_i. The matrices P and Q are *unitary* by construction, and the remaining pencil $\lambda\tilde{B}-\tilde{A}$ has no finite eigenvalues. Consequently, this last pencil has all its elementary divisors "at infinity." We now show how to proceed with this special point.

Let us therefore compare the canonical form of the elementary divisors at α and at infinity:

$$\begin{bmatrix} \lambda-\alpha & & & \\ -1 & \ddots & & \\ & \ddots & \ddots & \\ & & -1 & \lambda-\alpha \end{bmatrix}, \qquad \begin{bmatrix} -1 & & & \\ \lambda & \ddots & & \\ & \ddots & \ddots & \\ & & \lambda & -1 \end{bmatrix}. \tag{3.18a, b}$$

The structure of the first block (3.18a) is reconstructed by working on the expansion at α:

$$(\lambda-\alpha)B_{1,1}-A_{1,1}=(\lambda-\alpha)I-\begin{bmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix}.$$

The rank deficiency of $A_{1,1}$ (and of successive deflated $A_{j,j}$ blocks) will tell the nature of the elementary divisor at α. Analogously we could inspect the rank deficiency of $B_{1,1}$ in the expansion of the second block (3.18b),

$$\lambda B_{1,1}-A_{1,1}=\lambda\begin{bmatrix}0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0\end{bmatrix}-I,$$

in order to study its structure at infinity. This is done by following algorithm:

ALGORITHM 3.6

```
        comment initialization ;
        $j:=1;\ A_{1,1}:=A;\ B_{1,1}:=B;\ n_1:=n;$
step_j: comment compute the S.V.D. of the $n_j\times n_j$ matrix $B_{j,j}$ ;
        call SVD ($B_{j,j}$) result ($U_b,\Sigma_b,V_b$) nullity ($s_j$);
        if $s_j=0$ then begin $l:=j-1$; stop end;
        comment compress $B_{j,j}$ to full column rank and partition ;
        $\left[\,B_{j+1}\mid 0\,\right]:=B_{j,j}V_b;\ \left[\,A_{j+1}\mid A_j\,\right]:=A_{j,j}V_b;$
        comment also update and partition blocks in column $j$ ;
        for $i=1$ step 1 until $j-1$ do
          begin $\left[\,B_{j+1,i}\mid B_{j,i}\,\right]:=B_{j,i}V_b;\ \left[\,A_{j+1,i}\mid A_{j,i}\,\right]:=A_{j,i}V_b$ end;
        comment compute the S.V.D. of the $n_j\times s_j$ matrix $A_j$ ;
        call SVD ($A_j$) result ($U_a,\Sigma_a,V_a$) rank ($s_j$);
        comment compress $A_j$ to full row rank, permute and partition ;
        $\left[\frac{B_{j+1,j+1}}{B_{j+1,j}}\right]:=P_aU_a^*B_{j+1};\ \left[\frac{A_{j+1,j+1}}{A_{j+1,j}}\right]:=P_aU_a^*A_{j+1};\ \left[\frac{0}{A_{j,j}}\right]$
                    $:=P_aU_a^*A_j;$
        comment update ;
        $n_{j+1}:=n_j-s_j;\ j:=j+1;$ go to step _j ;
```

Analogously to (3.8), this algorithm results in

$$P(\lambda B-A)Q=\left[\begin{array}{c|c}\lambda B_{l+1,l+1}-A_{l+1,l+1} & 0\\ \hline X & \lambda B_{\infty}-A_{\infty}\end{array}\right]$$

$$=\left[\begin{array}{c|c|c|c|c}\lambda B_{l+1,l+1}-A_{l+1,l+1} & 0 & \cdots & 0 & 0\\ \hline X & -A_{l,l} & \cdots & 0 & 0\\ \hline \vdots & \vdots & & \vdots & \vdots\\ \hline X & \lambda B_{l,2}-A_{l,2} & \cdots & -A_{2,2} & 0\\ \hline \underbrace{\qquad X\qquad}_{n_{l+1}} & \underbrace{\lambda B_{l,1}-A_{l,1}}_{s_l} & \cdots & \underbrace{\lambda B_{2,1}-A_{2,1}}_{s_2} & \underbrace{-A_{1,1}}_{s_1}\end{array}\right]\begin{array}{l}\}n_{l+1}\\ \}s_l\\ \\ \}s_2\\ \}s_1\end{array}\tag{3.19}$$

where

(1) $B_{l+1,l+1}$ has full rank,
(2) the $A_{i,i}$ have full rank s_i $(i=1,\dots,l)$,
(3) the $B_{i,i-1}$ have full column rank s_i $(i=2,\dots,l)$.

PROPOSITION 3.7. *The indices $\{s_i\}$ given by Algorithm* 3.6 *completely determine the structure at infinity of the pencil $\lambda B - A$:*

$$\lambda B - A \textit{ has } (s_j - s_{j+1}) = d_j \textit{ elementary divisors of degree } j \qquad (j=1,\dots,l)$$

Proof. By using similar lemmas to 3.2, 3.3 and 3.4 it is easy to prove that $\lambda B - A$ is strictly equivalent to

$$\lambda B - A \sim \left[\begin{array}{c|cccc} \lambda B_{l+1,l+1} - A_{l+1,l+1} & & & & \\ \hline & -J_l & & & \\ & \lambda K_{l-1} & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda K_1 & -J_1 \end{array}\right] \tag{3.20}$$

with

$$J_i = I_{s_i} \quad \text{and} \quad K_i = \underbrace{\left[\begin{array}{c} I_{s_{i+1}} \\ \hline 0 \end{array}\right]}_{s_{i+1}} \left.\vphantom{\begin{array}{c} I \\ 0 \end{array}}\right\} s_i .$$

Here again a separation of structure elements is performed by the algorithm: the top pencil only has finite eigenvalues, since $B_{l+1,l+1}$ has full rank and the bottom pencil has all its eigenvalues at infinity. This last pencil is in quasi-Kronecker form and can be brought to canonical form by permutations only (this is carefully done for a more general case in Proposition 4.3), yielding the required result. ■

COMMENTS.

(a) In order to determine an elementary divisor, the algorithms described in this paper need an expansion in the associated eigenvalue. For finite eigenvalues, the expansion is

$$(\lambda-\alpha)B-(A-\alpha B)=\lambda' B_{1,1}-A_{1,1}, \tag{3.21}$$

and thus knowledge of α is required. For infinite eigenvalues, though, we can consider

$$\lambda B-A=\lambda B_{1,1}-A_{1,1} \tag{3.22}$$

as an expansion around $\lambda=\infty$, and no *a priori* knowledge is required (this can be interpreted as "infinity being always known").

Algorithms 3.1 and 3.6 use these expansions to determine the eigenstructure at the concerned point. This is done by inspecting the rank deficiency of (3.21) and (3.22) respectively. Some authors define the structure at ∞ as the structure at the eigenvalue zero of the transformed pencil $B_{1,1}-\lambda' A_{1,1}$ (implying the transformation $\lambda=1/\lambda'$ which maps $\lambda=\infty$ into $\lambda'=0$). This again illustrates the analogy between the two algorithms: their only difference is the interchange between A and B.

(b) When using Algorithm 3.6 on the deflated pencil $\lambda\tilde{B}-\tilde{A}$ of the form (3.17), one completes the structure analysis of the pencil $\lambda B-A$. This first pencil indeed has no finite eigenvalues and thus will be transformed to the staircase form $\lambda B_\infty - A_\infty$ yielding the structure at infinity. For numerical purposes it is preferable to start with the point at infinity. The recommended path is

(i) transform $\lambda B-A$ to the form

$$\lambda B-A\sim\left[\begin{array}{c|c}\lambda B_f-A_f & \\ \hline X & \lambda B_\infty-A_\infty\end{array}\right] \tag{3.23}$$

yielding the infinite elementary divisors and a deflated pencil λB_f-A_f with only finite eigenvalues.

(ii) separate from λB_f-A_f the Jordan structure of each eigenvalue, starting with the smallest one in norm.

By keeping this schedule the "shifts" $A-\alpha B$ are performed in increasing order, which guarantees better results [in step (i) no shifts are required]. At that stage, the Kronecker canonical form of $\lambda B-A$ is "recognizable" and only stable transformations have been used. If the transformation matrices that bring $\lambda B-A$ to its canonical form are needed, then one must proceed beyond the staircase form and perform transformations as described in Lemmas 3.2 and 3.3. These are Gaussian-type *eliminations without pivoting*. The computation of this information may thus be very *unstable*.

This problem arises when eigenvectors or eigenspaces related to an eigenvalue are needed. When dealing with the pencil $\lambda I - A$, Golub and Wilkinson [4] suggest restarting Kublanovskaya's algorithm on the original pencil for each eigenvalue, in order to obtain an orthogonal basis for each eigenspace independently. This demands a larger computational effort but avoids the unstable eliminations mentioned above. The same procedure can be applied to regular pencils using present algorithms.

(c) When the eigenvalues α_j are not known, one can still perform step (i) of (b) and determine the structure at infinity of $\lambda B - A$. In a following step, the QZ algorithm can be used on the deflated pencil $\lambda B_f - A_f$. This is a safer path than using QZ directly on the pencil $\lambda B - A$. Indeed, if infinite elementary divisors of degree higher than 1 occur in the pencil, they will not be recognized as such [1]. Worse, if their degrees are rather high, the corresponding infinite eigenvalues may be recognized numerically as finite ones.

EXAMPLE 2. Suppose $\lambda B - A$ has one finite eigenvalue at $\lambda = 20$ and one elementary divisor of degree 15 at infinity:

$$\lambda B - A = \lambda \left[\begin{array}{cccccc|c} 0 &&&&&& \\ 1 & \cdot &&&&& \\ & \cdot & \cdot &&&& \\ && \cdot & \cdot &&& \\ &&& \cdot & \cdot && \\ &&&& 1 & 0 & \\ \hline &&&&&& 1 \end{array}\right] - \left[\begin{array}{cccccc|c} 1 &&&&&& \\ & \cdot &&&&& \\ && \cdot &&&& \\ &&& \cdot &&& \\ &&&& \cdot && \\ &&&&& 1 & \\ \hline &&&&&& 20 \end{array}\right] \left.\vphantom{\begin{array}{c}1\\1\\1\\1\\1\\1\\1\end{array}}\right\} 16$$

Under perturbations of size ε (i.e. $\|E\|_2/\|B\|_2 \leqslant \varepsilon$ and $\|F\|_2/\|A\|_2 \leqslant \varepsilon$), the pencil $\lambda(B+E)-(A+F)$ may have perturbed eigenvalues as

$$|\lambda_i'| \cong \varepsilon^{-1/15} \qquad \text{for} \quad i = 1,\ldots,15$$

$$\lambda_{16}' = 20 + \eta \qquad \text{with} \quad |\eta| \leqslant 20\varepsilon.$$

On a computer with machine precision $\varepsilon = 10^{-15}$ the infinite eigenvalues may thus wander off severely and become in a certain sense more "finite" than the remaining eigenvalue $\lambda_{16} = 20$. This phenomenon is inevitable when computing the eigenstructure of $U(\lambda B - A)V$—where U and V are arbitrary unitary matrices—with the QZ algorithm. With Algorithm 3.6, on the other hand, we can first try to identify the structure of this pencil at ∞. Even when thereby only part of the structure is recognized, this is of considerable help for the problem.

TABLE 1

<table>
<tr><th>Algorithm 3.1</th><th>Dual Algorithm</th></tr>
<tr><td>$(\lambda-\alpha)B_{1,1}-A_{1,1}$</td><td>$(\lambda-\alpha)B_{1,1}-A_{1,1}$</td></tr>
<tr><td>$\Downarrow$</td><td>$\Downarrow$</td></tr>
<tr><td>*Step 1*</td><td>*Step 1*</td></tr>
<tr><td>$(\lambda-\alpha)[\,B_2 \mid B_1\,]-[\,A_2 \mid \underbrace{\leftarrow\,]}_{s_1}$</td><td>$\left.(\lambda-\alpha)\left[\begin{array}{c}\hat{B}_1\\ \hline B_{\hat{2}}\end{array}\right]-\left[\begin{array}{c}\downarrow\\ \hline \hat{A}_2\end{array}\right]\right\}\hat{s}_1$</td></tr>
<tr><td>$\Downarrow$</td><td>$\Downarrow$</td></tr>
<tr><td>$(\lambda-\alpha)\left[\begin{array}{c|c}B_{2,2}&\downarrow\\ \hline B_{2,1}&B_{1,1}\end{array}\right]-\left[\begin{array}{c|c}A_{2,2}&0\\ \hline A_{2,1}&\underbrace{0}_{s_1}\end{array}\right]\}s_1$</td><td>$(\lambda-\alpha)\left[\begin{array}{c|c}\hat{B}_{1,1}&\leftarrow\\ \hline \hat{B}_{2,1}&\hat{B}_{2,2}\end{array}\right]-\left[\begin{array}{c|c}0&0\\ \hline \underbrace{\hat{A}_{2,1}}_{\hat{s}_1}&\hat{A}_{2,2}\end{array}\right]\}\hat{s}_1$</td></tr>
</table>

(d) *Dual algorithm*. We can obtain dual forms for both Algorithms 3.1 and 3.6 by replacing column compressions with row compressions, and conversely, in each of these algorithms. This is pictured in Table 1 for Algorithm 3.1 and its dual form (compressions are indicated by arrows). In both cases the nullity of $A_{1,1}$ is computed and thus $s_1=\hat{s}_1$. In the first case s_1 eigenvalues at α are deflated at the bottom [as $(\lambda-\alpha)B_{1,1}$], while in the second case this is done at the top [as $(\lambda-\alpha)\hat{B}_{1,1}$]. The smaller pencils $(\lambda-\alpha)B_{2,2}-A_{2,2}$ and $(\lambda-\alpha)\hat{B}_{2,2}-\hat{A}_{2,2}$ are then processed as in the first step, etc. The sequence $\{\hat{s}_i\}$ can again be proven to yield the structure at α through the same formulas as for the $\{s_i\}$ sequence. Therefore $s_i=\hat{s}_i$, and thus the resulting staircase shape in each of the two algorithms is the pertranspose of the other. The same holds for Algorithm 3.6 and its dual form. Although these dual algorithms do not bring in anything new when applied to regular pencils, they play an important role in the singular case.

IV. SINGULAR PENCILS

As shown in Sec. II, the structure of singular pencils is more complex than that of regular pencils because of the occurrence of an additional structure element: the Kronecker indices. In this section we demonstrate how with minor changes in Algorithms 3.1 and 3.6 we can cope with singular pencils.

For better understanding, let us again look at blocks in canonical form and examine the action of Algorithm 3.6 on each of them. The structure elements are

$$\begin{bmatrix} \lambda-\alpha & & & & \\ -1 & \cdot & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & \cdot \\ & & & -1 & \lambda-\alpha \end{bmatrix}, \quad \begin{bmatrix} -1 & & & & \\ \lambda & \cdot & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & \cdot \\ & & & \lambda & -1 \end{bmatrix}, \tag{4.1a, b}$$

$$\begin{bmatrix} -1 & & & \\ \lambda & \cdot & & \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ & & & \cdot & -1 \\ & & & & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & -1 & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & \cdot \\ & & & \lambda & -1 \end{bmatrix}. \tag{4.1c, d}$$

Algorithm 3.6 reacts on pencils with defective *column* rank in $B_{1,1}$, the coefficient of λ in the expansion $\lambda B_{1,1}-A_{1,1}$. Thereby only (4.1b) and (4.1d) are detected. The dual algorithm reacts on the defective *row* rank of $B_{1,1}$ and therefore retrieves (4.1b) and (4.1c). Thus, in both cases column or row Kronecker indices are carried along with the infinite elementary divisors. Algorithm 3.1 looks for defective *column* (*row* for the dual algorithm) rank in $A_{1,1}$, the constant term in the expansion $(\lambda-\alpha)B_{1,1}-A_{1,1}$. Again this algorithm suffers from the same disease, since column (row) indices are carried along together with the finite elementary divisors at α.

These problems are to be expected, since both algorithms look for rank deficiency (in columns or rows) of a constant matrix, obtained from an expansion at the eigenvalue considered. For Kronecker blocks this matrix is nonsquare and thus rank-deficient (in columns or rows). (Some authors use this argument to say that *any* point in the complex plane is an eigenvalue of such a Kronecker block.) It appears now that previous algorithms, adapted for singular pencils, yield the required information *together* with the Kronecker indices. The only modification of Algorithm 3.6 required when dealing with singular pencils comes from the nonsquareness of the diagonal blocks $A_{j,j}$. The rank compression of the block A_j results in a $A_{j,j}$ of full *row* rank r_j, which may now be different from s_j.

ALGORITHM 4.1

```
comment initialization ;
$j:=1$; $A_{1,1}:=A$; $B_{1,1}:=B$; $m_1:=m$; $n_1:=n$;
```

```
step_j: comment compute the S.V.D. of the $m_j \times n_j$ matrix $B_{j,j}$ ;
    call SVD $(B_{j,j})$ result $(U_b, \Sigma_b, V_b)$ column nullity $(s_j)$;
    if $s_j = 0$ then begin $l := j - 1$; stop end;
    comment compress $B_{j,j}$ to full column rank and partition ;
    $[B_{j+1} \mid 0] := B_{j,j} V_b$; $[A_{j+1} \mid A_j] := A_{j,j} V_b$;
    comment also update and partition blocks in column $j$ ;
    for $i = 1$ step 1 until $j - 1$ do
        begin $[B_{j+1,i} \mid B_{j,i}] := B_{j,i} V_b$; $[A_{j+1,i} \mid A_{j,i}] := A_{j,i} V_b$ end ;
    comment compute the S.V.D. of the $m_j \times s_j$ matrix $A_j$ ;
    call SVD $(A_j)$ result $(U_a, \Sigma_a, V_a)$ rank $(r_j)$;
    comment compress $A_j$ to full row rank, permute and partition ;
    $\left[\begin{array}{c} B_{j+1,j+1} \\ \hline B_{j+1,j} \end{array}\right] := P_a U_a^* B_{j+1}$; $\left[\begin{array}{c} A_{j+1,j+1} \\ \hline A_{j+1,j} \end{array}\right] := P_a U_a^* A_{j+1}$; $\left[\begin{array}{c} 0 \\ \hline A_{j,j} \end{array}\right]$
    $:= P_a U_a^* A_j$;
    comment update ;
    $m_{j+1} := m_j - r_j$; $n_{j+1} := n_j - s_j$; $j := j + 1$; go to step _j
```

Note that the stopping rule of the algorithm is when $B_{j,j}$ has full *column* rank (then $s_j = 0$ and $l = j - 1$).

Hence we have constructed unitary matrices P and Q that reduce $\lambda B - A$ to the following form:

$$P(\lambda B - A)Q = \left[\begin{array}{c|c} \lambda B_{l+1,l+1} - A_{l+1,l+1} & 0 \\ \hline X & \lambda\tilde{B} - \tilde{A} \end{array}\right]$$

$$= \begin{array}{c} \left[\begin{array}{c|c|c|c|c} \lambda B_{l+1,l+1} - A_{l+1,l+1} & 0 & \cdots & 0 & 0 \\ \hline \lambda B_{l+1,l} - A_{l+1,l} & -A_{l,l} & \cdots & 0 & 0 \\ \hline \hline \lambda B_{l+1,2} - A_{l+1,2} & \lambda B_{l,2} - A_{l,2} & \cdots & -A_{2,2} & 0 \\ \hline \lambda B_{l+1,1} - A_{l+1,1} & \lambda B_{l,1} - A_{l,1} & \cdots & \lambda B_{2,1} - A_{2,1} & -A_{1,1} \end{array}\right] \begin{array}{l} \}\, m_{l+1} \\ \}\, r_l \\ \\ \}\, r_2 \\ \}\, r_1 \end{array} \\ \begin{array}{ccccc} \underbrace{\hphantom{\lambda B_{l+1,1} - A_{l+1,1}}}_{n_{l+1}} & \underbrace{\hphantom{\lambda B_{l,1} - A_{l,1}}}_{s_l} & \hphantom{\cdots} & \underbrace{\hphantom{\lambda B_{2,1} - A_{2,1}}}_{s_2} & \underbrace{\hphantom{-A_{1,1}}}_{s_1} \end{array} \end{array}, \tag{4.2}$$

where

(1) $B_{l+1,l+1}$ has full *column* rank,
(2) the $A_{i,i}$ have full *row* rank r_i $(i = 1, \ldots, l)$,
(3) the $B_{i,i-1}$ have full *column* rank s_i $(i = 2, \ldots, l)$.

From this it clearly follows that (putting $s_{l+1}=0$)

$$\begin{aligned} s_i - r_i \;&= e_i \geqslant 0 \quad \text{for} \quad i=1,\dots,l \\ r_i - s_{i+1} &= d_i \geqslant 0 \quad \text{for} \quad i=1,\dots,l \end{aligned}$$

LEMMA 4.2. *The pencil* $\lambda B - A$ *is strictly equivalent to the pencil* $\operatorname{diag}\{\lambda B_{l+1,l+1} - A_{l+1,l+1}, \lambda B_n - A_n\}$, *where*

$$\lambda B_n - A_n = \begin{bmatrix} -J_l & & & \\ \lambda K_{l-1} & \ddots & & \\ & \ddots & \ddots & \\ & & \lambda K_1 & -J_1 \end{bmatrix} \tag{4.3}$$

and

$$J_i = \underbrace{\left[\, I_{r_i} \mid 0 \,\right]}_{s_i}\Big\}r_i \quad \text{and} \quad K_i = \left[\begin{array}{c} I_{s_{i+1}} \\ \hline \underbrace{0}_{s_{i+1}} \end{array}\right]\Bigg\}r_i.$$

Proof. The proof is similar to those of Lemmas 3.2, 3.3 and 3.4. It is easy to check that in each of these steps the blocks $A_{i,i}$ considered only need to have full *row* rank (instead of being invertible). This is also the reason for the new definition of the blocks K_i and J_i. These extensions make use of the left or right invertibility of the blocks considered [see (2.3)]. ■

PROPOSITION 4.3. *The indices* $\{e_i\}$ *and* $\{d_i\}$ *given by Algorithm 4.1 completely determine the Kronecker column indices* $\{\varepsilon_i\}$ *and the infinite elementary divisors with their degrees* $\{\delta_i\}$.

Proof. Using Lemma 4.2 on the form (4.2), we have that $\lambda B - A$ is strictly equivalent to $\operatorname{diag}\{\lambda B_{l+1,l+1} - A_{l+1,l+1}, \lambda B_n - A_n\}$. Since $B_{l+1,l+1}$ has full column rank, the first pencil has no infinite elementary divisors and no Kronecker indices. Below, we show that by choosing appropriate permutations, $\lambda B_n - A_n$ consists simply and solely of these two structure elements. This is proved inductively:

In $\lambda B_n - A_n$ there is a λ under each -1 and a -1 to the right of each λ. This is because

$$0 = s_{l+1} \leqslant r_l \leqslant s_l \leqslant \cdots \leqslant r_2 \leqslant s_2 \leqslant r_1 \leqslant s_1. \tag{4.4}$$

From this pencil we extract

(1) d_l infinite elementary divisors of degree l,
(2) e_l blocks L_{l-1} with Kronecker column index $l-1$ as follows.

We start with the $d_l = r_l$ ones of J_l and separate, with permutations only (by extracting a 1 from each J_i and K_i) d_l blocks of the type

$$\left.\begin{bmatrix} -1 & & & \\ \lambda & \ddots & & \\ & \ddots & \ddots & \\ & & \lambda & -1 \end{bmatrix}\right\} l.$$

By doing this, each r_i and s_i in (4.4) decreases by r_l, and thus r_l becomes zero. Hence in λK_{l-1} there are $e_l = (s_l - r_l)$ λ's remaining. With these we analogously separate e_l blocks of the type

$$\left.\begin{bmatrix} \lambda & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & \ddots \\ & & \lambda & -1 \end{bmatrix}\right\} l-1.$$

After these two steps J_l and K_{l-1} have disappeared in $\lambda B_n - A_n$ and the remaining block sizes r_i and s_i are lowered by s_l. The inequalities (4.4) are preserved, as well as the differences $e_i = s_i - r_i$ and $d_i = r_i - s_{i+1}$. We thus have a pencil $\lambda B_n' - A_n'$ satisfying (4.4), but with l reduced by one. This completes the induction step. ∎

COROLLARY 4.4. *The indices* $\{e_i | i = 1,\dots,l\}$ *and* $\{d_i | i = 1,\dots,l\}$ *determine the following structure elements of the pencil* $\lambda B - A$:

(i) *there are* d_i *infinite elementary divisors of degree* i $(i = 1,\dots,l)$,
(ii) *there are* e_i *Kronecker blocks* L_{i-1} *of size* $i-1$ $(i = 1,\dots,l)$. ∎

EXAMPLE 3. Algorithm 4.1 applied to Example 1 transforms the matrix into the staircase form shown in Fig. 4. Here the bottom right matrix has already the form $\lambda B_n - A_n$ and the indices are $d_1 = 1$, $d_2 = 1$, $d_3 = 0$; $e_1 = 2$, $e_2 = 1$, $e_3 = 1$. By Corollary 4.4 we thus have $\delta_1 = 1$, $\delta_2 = 2$; $\varepsilon_1 = 0$, $\varepsilon_2 = 0$, $\varepsilon_3 = 1$, $\varepsilon_4 = 2$.

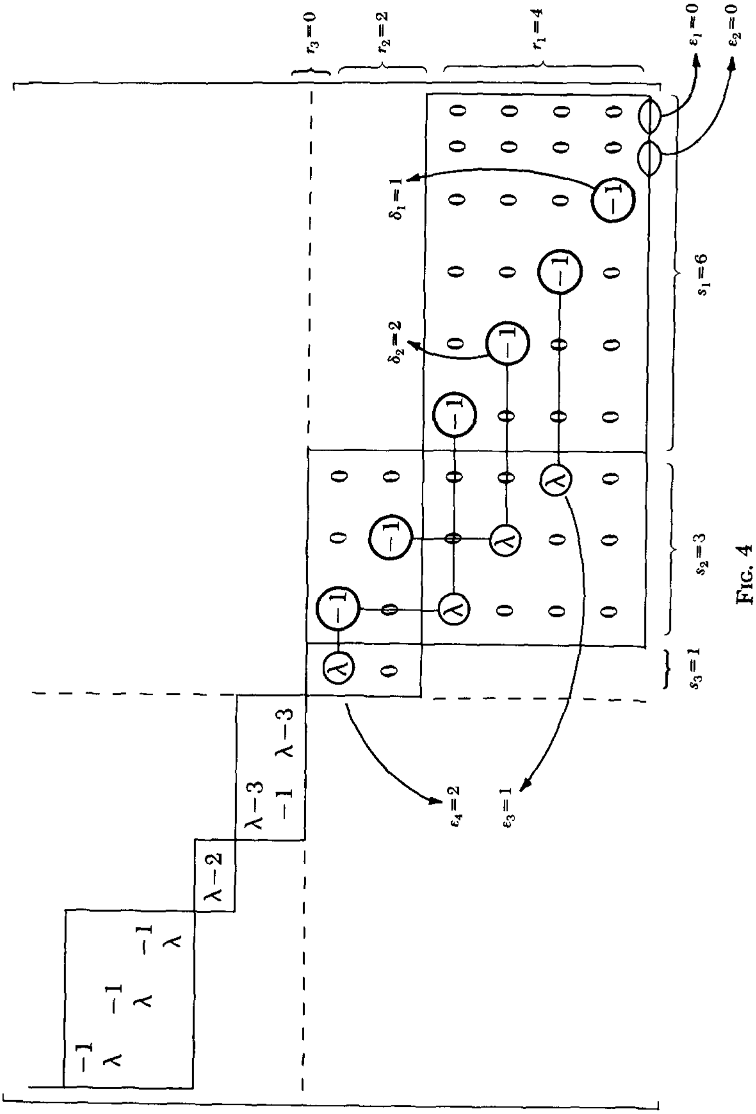
$r_3=0$
$r_2=2$
$r_1=4$
$\varepsilon_1=0$
$\varepsilon_2=0$
$\delta_1=1$
$\delta_2=2$
$s_1=6$
$s_2=3$
$s_3=1$
$\varepsilon_4=2$
$\varepsilon_3=1$
λ
-1
0
$\lambda-2$
$\lambda-3$

FIG. 4

The recovery of the original blocks in $\lambda B_n - A_n$, as explained in Proposition 4.3, is also indicated in Fig. 4.

The idea of the algorithm is easily seen in this example. Here we have separated recursively the structure elements whose coefficient of λ has defective *column* rank from the others. The dual algorithm has a similar effect by acting on the *row* rank of the coefficient of λ. Thereby the *row* Kronecker indices are detected together with the structure at infinity [the blocks (4.1c) and (4.1b) indeed have deficient *row* rank in the coefficient of λ]. The duality of the algorithm thus lies in the interchange of row and column compressions:

ALGORITHM 4.5.

```
        comment initialization;
        $j := 1;\ A_{1,1} := A;\ B_{1,1} := B;\ m_1 := m;\ n_1 := n;$
step_j: comment compute the S.V.D. of the $m_j \times n_j$ matrix $B_{j,j}$ ;
        call SVD $(B_{j,j})$ result $(U_b, \Sigma_b, V_b)$ nullity $(\hat{s}_j)$;
        if $\hat{s}_j = 0$ then begin $k := j-1$; stop end;
        comment compress $B_{j,j}$ to full row rank, permute and partition ;
        $\left[\begin{array}{c} 0 \\ \hline B_{j+1} \end{array}\right] := P_b U_b^* B_{j,j};\ \left[\begin{array}{c} A_j \\ \hline A_{j+1} \end{array}\right] := P_b U_b^* A_{j,j};$
        comment also update and partition blocks in row $j$ ;
        for $i = 1$ step 1 until $j-1$ do
            begin $\left[\begin{array}{c} B_{j,i} \\ \hline B_{j+1,i} \end{array}\right] := P_b U_b^* B_{j,i};\ \left[\begin{array}{c} A_{j,i} \\ \hline A_{j+1,i} \end{array}\right] := P_b U_b^* A_{j,i}$ end;
        comment compute the S.V.D. of the $\hat{s}_j \times n_j$ matrix $A_j$ ;
        call SVD $(A_j)$ result $(U_a, \Sigma_a, V_a)$ rank $(\hat{r}_j)$;
        comment compress $A_j$ to full column rank and partition ;
        $\left[\begin{array}{c|c} B_{j+1,j} & B_{j+1,j+1} \end{array}\right] := B_{j+1} V_a;\ \left[\begin{array}{c|c} A_{j+1,j} & A_{j+1,j+1} \end{array}\right] := A_{j+1} V_a;$
        $\left[\begin{array}{c|c} A_{j,j} & 0 \end{array}\right] := A_j V_a;$
        comment update; $m_{j+1} := m_j - \hat{s}_j;\ n_{j+1} := nj - \hat{r}_j;\ j := j+1;$ go to
        step _j;
```

This "dual" algorithm reduces $\lambda B - A$ to the following form (the staircase shape is "pertransposed" to the one resulting from the actual algorithm):

$$P(\lambda B - A)Q = \left[\begin{array}{c|c} \lambda\hat{B} - \hat{A} & 0 \\ \hline X & \lambda B_{k+1,k+1} - A_{k+1,k+1} \end{array}\right]$$

$$
=\left[\begin{array}{c|c|c|c|c}
-A_{1,1} & 0 & \cdots & 0 & 0 \\ \hline
\lambda B_{2,1}-A_{2,1} & -A_{2,2} & \cdots & 0 & 0 \\ \hline
\vdots & \vdots & & \vdots & \vdots \\ \hline
\lambda B_{k,1}-A_{k,1} & \lambda B_{k,2}-A_{k,2} & \cdots & -A_{k,k} & 0 \\ \hline
\underbrace{X}_{\hat{r}_1} & \underbrace{X}_{\hat{r}_2} & \cdots & \underbrace{X}_{\hat{r}_k} & \underbrace{\lambda B_{k+1,k+1}-A_{k+1,k+1}}_{n_{k+1}}
\end{array}\right]
\begin{array}{l}
\}\,\hat{s}_1 \\ \}\,\hat{s}_2 \\ \\ \}\,\hat{s}_k \\ \}\,m_{k+1}
\end{array},
\tag{4.5}
$$

where

(1) $B_{k+1,k+1}$ has full row rank,
(2) the $A_{i,i}$ blocks have full column rank $\hat{r}_i$,
(3) the $B_{i,i-1}$ blocks have full row rank $\hat{s}_i$.

Hence relying on the duality between both results, we have

COROLLARY 4.6. *Algorithm 4.5 applied to a singular pencil $\lambda B-A$ determines its Kronecker row indices $\{\eta_i\}$ and infinite elementary divisors $\{\delta_i\}$ as follows:*

(i) *there are $\hat{d}_i=(\hat{r}_i-\hat{s}_{i+1})$ infinite elementary divisors of degree i $(i=1,\dots,k)$*

(ii) *there are $\hat{e}_i=(\hat{s}_i-\hat{r}_i)$ Kronecker blocks L_{i-1}^{P} of size $i-1$ $(i=1,\dots,k)$*

Algorithms 4.1 and 4.5 together yield the following result.

PROPOSITION 4.7. *It is always possible to construct unitary transformations that put an arbitrary pencil in the following form:*

$$
P(\lambda B-A)Q=\left[\begin{array}{c|c|c|c}
\lambda B_{\eta}-A_{\eta} & & & \\ \hline
X & \lambda B_f-A_f & & \\ \hline
X & X & \lambda B_{\infty}-A_{\infty} & \\ \hline
X & X & X & \lambda B_{\varepsilon}-A_{\varepsilon}
\end{array}\right],
\tag{4.6}
$$

where

(i) *λB_f-A_f is a square regular pencil containing the finite elementary divisors of $\lambda B-A$.*

(ii) $\lambda B_{\infty}-A_{\infty}$ *is a square regular pencil containing the infinite elementary divisors of* $\lambda B-A$.

(iii) $\lambda B_{\eta}-A_{\eta}$ *and* $\lambda B_{\varepsilon}-A_{\varepsilon}$ *are singular pencils containing the Kronecker row and column structure respectively.*

Proof. Observe that $\lambda B-A$ is unitarily equivalent to (4.2). In this pencil the top block $\lambda B_{l+1,l+1}-A_{l+1,l+1}$ has only two structure elements: Kronecker row indices and finite elementary divisors. Only the first of these two elements has defective *row* rank in the coefficient of λ [see (4.1c) and (4.1a)] and will thus be detected by Algorithm 4.5. Applying this algorithm thus yields the separation of these two structure elements:

$$P_1(\lambda B_{l+1,l+1}-A_{l+1,l+1})Q_1=\left[\begin{array}{c|c}\lambda B_{\eta}-A_{\eta} & \\ \hline X & \lambda B_f-A_f\end{array}\right]. \tag{4.7}$$

Similarly $\lambda\tilde{B}-\tilde{A}$ is transformed by Algorithm 4.5 to

$$P_2(\lambda\tilde{B}-\tilde{A})Q_2=\left[\begin{array}{c|c}\lambda B_{\infty}-A_{\infty} & \\ \hline X & \lambda B_{\varepsilon}-A_{\varepsilon}\end{array}\right], \tag{4.8}$$

in which the infinite elementary divisors are separated from the remaining $\lambda B_{\varepsilon}-A_{\varepsilon}$. Inserting (4.7) and (4.8) in (4.2), we obtain the desired result. ■

REMARKS.

(a) Algorithms 4.1 and 4.5 do not only perform the separation between the four structure elements, as shown in (4.6), but they also yield some internal information. Indeed, in (4.7) $\lambda B_{\eta}-A_{\eta}$ will have a specific shape yielding the Kronecker row indices. The "residual" λB_f-A_f has no special feature other than the invertibility of B_f. Since $\lambda\tilde{B}-\tilde{A}$ already gives the Kronecker column indices and the infinite elementary divisors, the reduction (4.8) can be omitted. At that stage, the only structure element to be recovered is the Jordan information of the pencil λB_f-A_f. This problem is discussed in a previous section.

(b) On the other hand, when a decomposition as in (4.6) is indeed required, the reduction (4.8) of $\lambda\tilde{B}-\tilde{A}$ must also be performed. When using Algorithm 4.5 for that purpose, the extracted part $\lambda B_{\infty}-A_{\infty}$ is maintained in staircase form, but the residual $\lambda B_{\varepsilon}-A_{\varepsilon}$ is not. Algorithm 4.1 can then be used again on $\lambda B_{\varepsilon}-A_{\varepsilon}$ in order to "rebuild" its staircase. It is possible to avoid this by adapting Algorithm 4.5. This is omitted here in the interests of brevity.

(c) When starting with Algorithm 4.5 on $\lambda B - A$ and next applying Algorithm 4.1 to the pencils $\lambda \hat{B} - \hat{A}$ and $\lambda B_{k+1,k+1} - A_{k+1,k+1}$ [see (4.5)], we obtain similarly

$$\hat{P}(\lambda B - A)\hat{Q} = \left[\begin{array}{c|c|c|c} \lambda B_\eta - A_\eta & & & \\ \hline X & \lambda B_\infty - A_\infty & & \\ \hline X & X & \lambda B_f - A_f & \\ \hline X & X & X & \lambda B_\varepsilon - A_\varepsilon \end{array}\right], \qquad (4.9)$$

whereby the regular pencils $\lambda B_\infty - A_\infty$ and $\lambda B_f - A_f$ still remain in the middle but are interchanged.

(d) The remarks above yield that the form (4.6) [or (4.9)] can be obtained by *unitary* equivalence transformations only, and that thereby one can ask for all the structure blocks to be in their staircase form. This reflects the numerical (backward) stability of the proposed algorithms for the determination of the structure of an arbitrary pencil.

(e) Algorithm 3.1 and its dual form can both be adapted to singular pencils also. This again yields algorithms to compute the Kronecker structure (*row* and *column*) together with finite elementary divisors at the considered point α. Hence when an expansion is used at a point which is *not* an eigenvalue of the pencil, these algorithms only extract the Kronecker column or row structure. They are not given here, in the interests of brevity. Moreover, from a numerical point of view, Algorithms 4.1 and 4.5 should be preferred because of the absence of any shift α.

V. APPLICATIONS

a.

The eigenvalue problem of a square pencil $\lambda B - A$ is currently regarded as "ill posed" when $\det(\lambda B - A) \equiv 0$ [6, 7]. The present algorithms, though, show that it is always possible (even in the nonsquare case) to reduce such pencils to the form (4.6) and thus extract a "regular part" $\lambda B_r - A_r$ with nonvanishing determinant. This regular part contains, in our language, the finite and infinite elementary divisors of the pencil.

We do not want to emphasize thereby that this part of the problem is well conditioned. On the contrary, it is easy to show that almost any perturbation of a *square singular pencil* will turn this pencil to a regular one. The Kronecker indices thus disappear and induce additional eigenvalues which can be *anywhere* in the complex plane, except that for real pencils the complex eigenvalues will of course be paired. Hence, *generically* (i.e. for random coefficient matrices) a square pencil is regular.

EXAMPLE 4. The following pencil has Kronecker indices $\varepsilon_1 = 1$ and $\eta_1 = 2$:

$$\begin{bmatrix} \lambda & -1 & & \\ & & -1 & \\ & & \lambda & -1 \\ & & & \lambda \end{bmatrix}. \tag{5.1}$$

A possible perturbation of this pencil is e.g. (with $|\varepsilon_i|$ smaller than ε, the machine precision)

$$\begin{bmatrix} \lambda & -1 & & \\ & \varepsilon_4\lambda + \varepsilon_3 & -1 & \\ & \varepsilon_2 & \lambda & -1 \\ \varepsilon_0 & \varepsilon_1 & & \lambda \end{bmatrix}. \tag{5.2}$$

Its determinant is $\varepsilon_4\lambda^4 + \varepsilon_3\lambda^3 + \varepsilon_2\lambda^2 + \varepsilon_1\lambda + \varepsilon_0$. Up to a scalar factor we can thus construct any polynomial and also any characteristic roots (infinity also, by choosing $\varepsilon_4 = 0$).

When applying the QZ algorithm to this pencil (or to any unitarily equivalent pencil), we would end up with a pencil (where x represents an arbitrary complex number) [6]

$$\lambda \begin{bmatrix} b_1 & x & x & x \\ & b_2 & x & x \\ & & b_3 & x \\ & & & b_4 \end{bmatrix} - \begin{bmatrix} a_1 & x & x & x \\ & a_2 & x & x \\ & & a_3 & x \\ & & & a_4 \end{bmatrix}, \tag{5.3}$$

where one of the ratios a_i/b_i has both a_i and b_i close to 0 and is thus ill posed. This signifies [7] that the pencil considered is very near to a singular one (e.g. because a_i and b_i are equal to 0). From then on, all other ratios should be distrusted, since they may correspond to "fake" eigenvalues. In the pencil (5.1) there are indeed three such eigenvalues $\lambda = 0$. These "fake" eigenvalues belong to Kronecker blocks with size larger than zero (see Example 4). Moreover, when a square singular pencil has a regular part, the QZ algorithm is not able to distinguish "regular" eigenvalues from "fake" ones. Worse, it is possible that none of the ratios corresponds to a "regular" eigenvalue, as shown in this example:

Example 5. The pencil

$$\left[\begin{array}{cccc}
\cline{1-2}
\multicolumn{1}{|c}{\lambda} & \multicolumn{1}{c|}{-1} & & \\
\cline{1-3}
 & & \multicolumn{1}{|c|}{(\lambda-2)} & \\
\cline{3-4}
 & & & \multicolumn{1}{|c|}{-1} \\
 & & & \multicolumn{1}{|c|}{\lambda} \\
\cline{4-4}
\end{array}\right] \tag{5.4}$$

is in the form (5.3) and has apparent eigenvalues 0/1, 0/0, 0/0 and 0/1.

Another drawback of the QZ algorithm is that the pencil treated can be close to a singular one even without the occurrence of such ill-conditioned ratios.

Example 6. The pencil

$$\lambda\begin{bmatrix} \varepsilon^{1/4} & -1 & & \\ & \varepsilon^{1/4} & -1 & \\ & & \varepsilon^{1/4} & -1 \\ & & & \varepsilon^{1/4} \end{bmatrix} - \begin{bmatrix} 0 & \varepsilon^{1/4} & -1 & \\ & 0 & \varepsilon^{1/4} & -1 \\ & & 0 & \varepsilon^{1/4} \\ & & & 0 \end{bmatrix} \tag{5.5}$$

has four eigenvalues $\lambda = 0$, and the diagonal elements b_i are all considerably larger than ε, the machine precision (take $\varepsilon = 10^{-16}$; then $\varepsilon^{1/4} = 10^{-4}$). Yet, the perturbed pencil

$$\lambda\begin{bmatrix} \varepsilon^{1/4} & -1 & & \\ & \varepsilon^{1/4} & -1 & \\ & & \varepsilon^{1/4} & -1 \\ -\varepsilon & & & \varepsilon^{1/4} \end{bmatrix} - \begin{bmatrix} 0 & \varepsilon^{1/4} & -1 & \\ & 0 & \varepsilon^{1/4} & -1 \\ & & 0 & \varepsilon^{1/4} \\ & -\varepsilon & & 0 \end{bmatrix}$$

is strictly equivalent (by a row transformation) to

$$\lambda\begin{bmatrix}\boxed{\begin{matrix}\varepsilon^{1/4} & -1 & & \\ & \varepsilon^{1/4} & -1 & \\ & & \varepsilon^{1/4} & -1\end{matrix}}\\ \\ \end{bmatrix}-\begin{bmatrix}\boxed{\begin{matrix}0 & \varepsilon^{1/4} & -1 & \\ & 0 & \varepsilon^{1/4} & -1 \\ & & 0 & \varepsilon^{1/4}\end{matrix}}\\ \\ \end{bmatrix}, \qquad (5.6)$$

which has Kronecker indices $\varepsilon_1=3$ and $\eta_1=0$.

We believe that the algorithms presented here can be of considerable help when dealing with such examples. Indeed, Algorithms 4.1 and 4.5 would at least detect the existence of a Kronecker structure when working on the pencils (5.1), (5.4) and (5.6) or on any ε-perturbation of each of these pencils. On the other hand, the algorithms will not necessarily retrieve the exact Kronecker indices, because the algorithms are based on the recognition of the rank deficiency of certain blocks. The singular-value decomposition, used for this purpose, is known to be a reliable tool, since even when the original matrix is not singular, this algorithm tells the user if there is any singular matrix in an ε-neighborhood of the original one. After (say) $i-1$ steps of our algorithm, though, rounding errors induced in previous steps may affect seriously the left-over pencil $\lambda B_{i,i}-A_{i,i}$. The rank search of this part thus relies on previous computations. Experiments have shown that this effect is amplified when the "rank-carrying stairs" (by which we mean the diagonal and off-diagonal blocks with full row or column rank in our staircase form) have small singular values. In (5.1) e.g. these blocks are all 1 and rounding errors will have no consequences, while in (5.6) some are $\varepsilon^{1/4}$ and the index $\varepsilon_1=3$ will not be retrieved. A detailed investigation of this phenomenon is in progress.

In the *rectangular case* the problem is very similar: for almost any perturbation most of the Kronecker structure disappears. This results from the fact that, generically, a nonsquare constant matrix has rank equal to its smallest dimension. Hence, when running Algorithm 4.5 on a "random" $m\times n$ pencil (with $m>n$) the following results are obtained. Let $\alpha=m-n$, and let the integers β and k be defined by $n=(k-2)\alpha+\beta$ (with $0\leqslant\beta<\alpha$). Then all constructed stairs will be square and of size α ($\hat{s}_i=\hat{r}_i=\alpha$) except for the last two stairs, for which $\hat{s}_{k-1}=\alpha$, $\hat{r}_{k-1}=\beta$ and $\hat{s}_k=\beta$, $\hat{r}_k=0$. The pencil will be completely processed after this. The structure of such a pencil then consists of:

$\hat{s}_{k-1}-\hat{r}_{k-1}=\alpha-\beta$ row indices equal to $k-2$,
$\hat{s}_k-\hat{r}_k \quad =\beta$ row indices equal to $k-1$.

A similar result holds for the case $m<n$. Hence, *generically* nonsquare

pencils have either only column indices (when $m<n$) or only row indices (when $m>n$), and these indices are determined by the dimensions of the pencil only. Results in this vein were also obtained by Wilkinson [23] and Wonham [24].

b.

Recently [17–19] attention has been given to the eigenvalue problem associated with the polynomial matrix

$$P(\lambda)=P_o+P_1\lambda+\cdots+P_l\lambda^l. \tag{5.7}$$

One common procedure is to "linearize" this matrix to the pencil

$$\lambda\begin{bmatrix} & I & & \\ & & \ddots & \\ & & & I \\ P_l & \cdots & \cdots & P_1 \end{bmatrix}-\begin{bmatrix} I & & & \\ & \ddots & & \\ & & I & \\ & & & -P_o \end{bmatrix}=\lambda S-T. \tag{5.8}$$

This pencil is regular when $P(\lambda)$ is regular $[\det P(\lambda)\neq 0]$ and has the same finite elementary divisors as $P(\lambda)$ [19], though the infinity structure can be quite different [16]. The current theory and algorithms for pencils can be used to reduce $\lambda S-T$ to a "minimal" pencil having the same structure as $P(\lambda)$ [16]. This procedure also applies to arbitrary polynomial matrices and produces a "minimal" singular pencil with the same structure as $P(\lambda)$. These principles are based on the Smith-Macmillan degree theory [16, 14].

c.

A well-known problem in systems theory is the *realization* of the "polar section" [20]:

$$R_\alpha(\lambda)=\frac{R_{-l}}{(\lambda-\alpha)^l}+\cdots+\frac{R_{-2}}{(\lambda-\alpha)^2}+\frac{R_{-1}}{(\lambda-\alpha)}. \tag{5.9}$$

The problem is to find minimal constant matrices $\{A_\alpha,B_\alpha,C_\alpha\}$ such that

$$R_\alpha(\lambda)=C_\alpha(\lambda I-A_\alpha)^{-1}B_\alpha. \tag{5.10}$$

Here the Jordan structure of A_2 reflects the polar structure of $R_\alpha(\lambda)$. A nonminimal solution is given by [12]:

$$A=\begin{bmatrix} \alpha I & I & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & I \\ & & & & \alpha I \end{bmatrix}, \qquad B=\begin{bmatrix} 0 \\ \vdots \\ 0 \\ I \end{bmatrix},$$

$$C=\begin{bmatrix} R_{-l} & \cdots & \cdots & R_{-1} \end{bmatrix}. \tag{5.11}$$

The extraction of a minimal realization from (5.11) can be viewed as the extraction of the Jordan structure at α from the pencil [20]:

$$\lambda\begin{bmatrix} I \\ 0 \end{bmatrix}-\begin{bmatrix} A \\ -C \end{bmatrix}.$$

Our algorithms provide an efficient method for performing this extraction.

d.

A combination of the previous two applications leads to the construction of a singular pencil having the same structure (i.e. finite structure, infinite structure and Kronecker indices) as a general rational matrix [16]. This can be viewed as a valid generalization of the eigenvalue problem to rational matrices. A perturbation analysis for rational matrices can e.g. easily be done through the study of this "strict equivalent pencil."

VI. CONCLUSION

In this paper we have presented a set of algorithms for the computation of the Kronecker canonical form of an arbitrary pencil. By exclusively using unitary transformations, the backward stability of the algorithms is guaranteed. Under such transformations we can obtain a "Schur type" decomposition (see Proposition 4.7) revealing its Kronecker structure. Since the determination of the Kronecker canonical form of a pencil is an ill-conditioned problem, the backward stability of our algorithms does *not* guarantee that the computed structure indeed corresponds to the pencil that was originally processed: any small error may affect this structure.

In each step of the algorithms we make in fact a specific choice: elements of smaller size than the machine accuracy are disregarded in order to obtain minimum rank for each stair. This fixes part of the Kronecker structure we are evaluating. While, in general, there is really little justification for such a choice, it makes sense in several applications. In the applications of Sec. V.b and c, e.g., this results in lower-order models for a given transfer function [16, 20]. Such small corrections are needed when one wants a "nongenerical result" such as a multiple eigenvalue at $\lambda = \infty$ (Sec. V.b) or any eigenvalue in the nonsquare case (Sec. V.c). For such applications, the problem could be formulated as an optimization problem: one is looking for a structure which is as degenerate (nongenerical) as possible. A similar approach to the classical eigenvalue problem has been shown to improve the conditioning of the problem [22]. Such policies again can only be justified through their use in certain applications.

Although the picture is still far from complete, we believe that the approach chosen in this paper will be helpful for understanding the nature of singular pencils.

The author is grateful to Professor P. Dewilde of the University of Delft for supervising the development of this work. Thanks are also due to Professor G. Golub and Professor J. H. Wilkinson for their helpful comments during the author's visit at Stanford University.

REFERENCES

1 J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon, Oxford, 1965.

2 F. R. Gantmacher, *Theory of Matrices*, Vols. I and II, Chelsea, New York, 1959.

3 J. H. Wilkinson and C. Reinsch, *Handbook for Automatic Computation, Vol. II: Linear Algebra*, Springer, New York, 1971.

4 G. H. Golub and J. H. Wilkinson, Ill-conditioned eigensystem and the computation of the Jordan canonical form, *SIAM Rev.* 18:578–619 (Oct. 1976).

5 V. N. Kublanovskaya, On a method of solving the complete eigenvalue problem for a degenerate matrix, *USSR Computational Math. and Math. Phys.* 6:1–14 (1968).

6 C. B. Moler and G. W. Stewart, An algorithm for the generalized matrix eigenvalue problem $Ax = \lambda Bx$, *SIAM J. Numer. Anal.* 10:241–256 (Apr. 1973.)

7 R. C. Ward, The combination shift QZ algorithm, *SIAM J. Numer. Anal.* 12:835–853 (Dec. 1975).

8 L. Kronecker, Algebraische Reduction der Schaaren bilinearer Formen, S. -*B. Akad.*, Berlin, 1890, pp. 763–776.

9 J. S. Thorp, An algorithm for an invariant canonical form, in *Proceedings of the 1974 IEEE Conference on Decision and Control*, pp. 248–253.

10 D. Jordan and L. F. Godbout, On the computation of the canonical pencil of a linear system. *IEEE Trans. Automatic Control* AC-22:112–114 (Feb. 1977).

11 J. S. Thorp, The singular pencil of a linear dynamical system, *Internat. J. Control* 18:577–596 (1973).
12 P. Van Dooren and P. Dewilde, State-space realization of a general rational matrix. A numerically stable algorithm, in *Proceedings of the* 1977 *Midwest Symposium on Circuits and Systems*, pp. 773–781.
13 A. S. Morse, Structural invariants of linear multivariable systems, *SIAM J. Control* 11:446–465 Aug. 1973.
14 H. H. Rosenbrock, *State-Space and Multivariable Theory*, Nelson, London, 1970.
15 B. Kouvaritakis, A geometric approach to the inversion of multivariable systems, *Internat. J. Control* 24:609–626 (1976).
16 P. Van Dooren, G. Verghese and T. Kailath, Properties of the system matrix of a generalized state-space system, in *Proceedings of the* 1978 *IEEE Conference on Decision and Control*, pp. 173–175.
17 A. Ruhe, Algorithms for the nonlinear eigenvalue problem, *SIAM J. Numer. Anal.* 10:694–689 (Sept. 1973).
18 G. Peters and J. H. Wilkinson, $Ax=\lambda BX$ and the generalized eigenproblem, *SIAM J. Numer. Anal.* 7:479–492 (Dec. 1970).
19 I. Gohberg, P. Lancaster and L. Rodman, Spectral analysis of matrix polynomials, Res. Paper No. 313, Univ. of Calgary, Calgary, 1976.
20 P. Van Dooren and P. Dewilde, Polar structure of rational matrices and the realization problem, Internal Report Div. Appl. Math Comput. Sci., K. Univ. Leuven, Louvain, 1977.
21 J. H. Wilkinson, The differential system $B\dot{x}=Ax$ and the generalized eigenvalue problem $Au=\lambda Bu$, Nat. Phys. Lab. Report NAC 73, Jan. 1977.
22 W. Kahan, Conserving confluence curbs ill-condition, Internal Report Dept. Comput. Sci., Univ. of Calif., Berkeley, 1972.
23 J. H. Wilkinson, private communication, 1978.
24 W. M. Wonham, *Linear Multivariable Theory. A Geometric Approach*, Springer, Berlin, 1974.

Received 14 February 1978; revised 5 December 1978