Lawrence Berkeley National Laboratory
Recent Work

Title

The Dimension of Matrices (Matrix Pencils) with Given Jordan (Kronecker) Canonical Forms

Permalink

Authors

Dummel, J.W.
Edelman, A.

Publication Date

1992

Lawrence Berkeley Laboratory
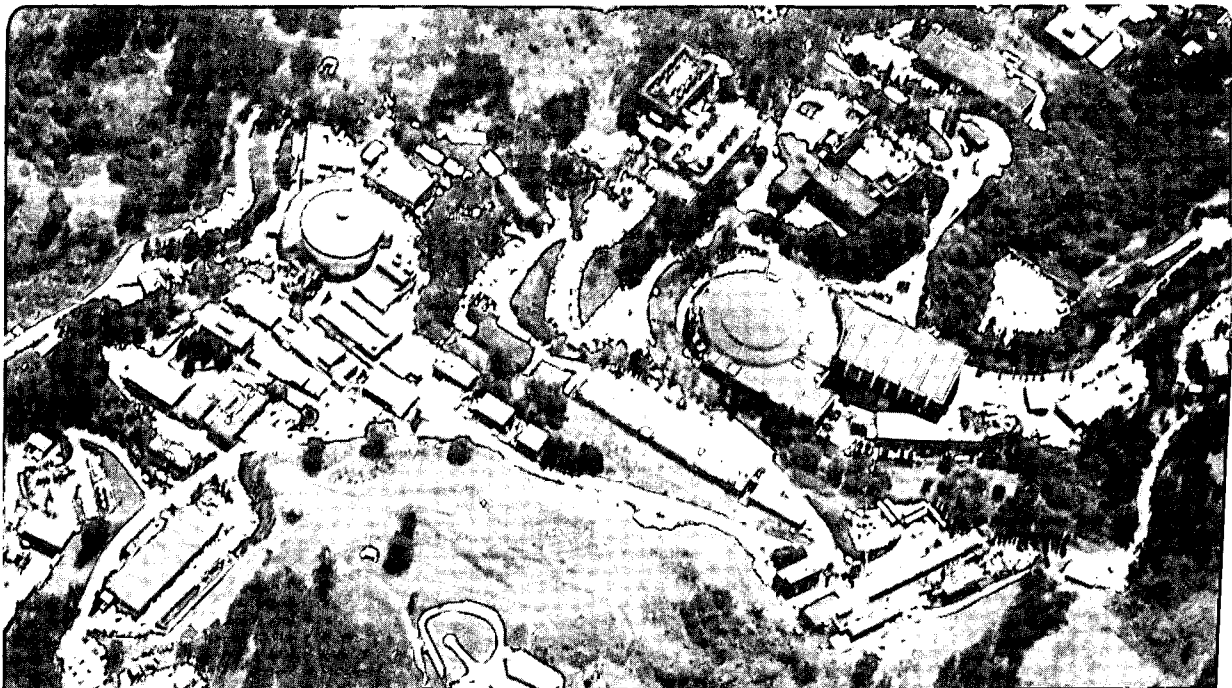## UNIVERSITY OF CALIFORNIA

Physics Division

Mathematics Department

To be submitted for publication

**The Dimension of Matrices (Matrix Pencils) with
Given Jordan (Kronecker) Canonical Forms**

J.W. Demmel and A. Edelman

January 1992

DISCLAIMER

DISCLAIMER

THE DIMENSION OF MATRICES (MATRIX PENCILS) WITH GIVEN JORDAN (KRONECKER) CANONICAL FORMS[1]

James W. Demmel[2]
Computer Science Division
and
Department of Mathematics
University of California
Berkeley, CA 94720

Alan Edelman[3]
Lawrence Berkeley Laboratory
and
Department of Mathematics
University of California
Berkeley, CA 94720

January 1992

The Dimension of Matrices (Matrix Pencils) with Given Jordan (Kronecker) Canonical Forms

James W. Demmel *
Alan Edelman†
University of California
Berkeley, California 94720

January 1992

Abstract

The set of n by n matrices with a given Jordan canonical form defines a subset of matrices in n^2 dimensional space. We analyze one classical approach and one new approach to count the dimension of this set. The new approach is based upon and meant to give insight into the staircase algorithm for the computation of the Jordan Canonical Form as well as the failures of these algorithms. We extend both techniques to count the dimension of the more complicated situation concerning the Kronecker canonical form of an arbitrary rectangular matrix pencil $A - \lambda B$.

1 Introduction

Given any square matrix A, the set of matrices similar to A forms a manifold in n^2 dimensional space. This manifold is, of course, the orbit of A under the action of conjugation:

$$\operatorname{orbit}(A) = \{ PAP^{-1} : \det(P) \neq 0 \},$$

The matrix pencil analog is to consider any pair of m by n matrices A and B, and define the orbit of the matrix pencil $A - \lambda B$ by the action of multiplication on the left and right by square nonsingular matrices of the appropriate size:

$$\text{orbit}(A - \lambda B) = \{P(A - \lambda B)Q^{-1} : \det(P)\det(Q) \neq 0\},$$

This orbit defines a manifold of pencils in $2mn$ dimensional space. All pencils on this manifold are said to be *equivalent* to $A - \lambda B$.

Our concern in this work is to count the (co)dimension of these manifolds. For simplicity of exposition, we sometimes refer to these two problems as counting the (co)dimension of a (single) matrix or of a matrix pencil, when more properly, we would refer to counting the (co)dimension of the orbits. We take two approaches, one based on classical techniques that identify the tangent spaces of these manifolds and the other based upon the algorithms as they exist [6, 7, 8, 9, 11, 12, 14, 15].

The classical approach to solving this problem requires the computation of the tangent space to the orbits. In the single matrix case, the tangent vectors have the form

$$XA - AX, \tag{1}$$

while in the matrix pencil case, the tangents have the form

$$X(A - \lambda B) - (A - \lambda B)Y. \tag{2}$$

Thus the codimension of the single matrix orbit is the number of linearly independent matrices X for which (1) vanishes, while the codimension of the matrix pencil orbit is related to the number of linearly independent matrix pairs X, Y for which (2) vanishes.

The new approach is based on the so called **staircase algorithms** for the Jordan and Kronecker canonical forms. The staircase algorithm for the Jordan canonical form proceeds by computing the Weyr characteristics of the matrix, while the staircase canonical form proceeds by computing a more complicated set of structural indices.

This new approach adds to our knowledge of the staircase algorithms for the canonical forms. In particular, it gives us geometric information regarding which manifold we are in at each stage of the algorithm.

Arnold [1] has rederived the formula for the Jordan case for the purpose of defining a particular normal form for deformations of a matrix with a given Jordan form. This form is convenient because of its minimum number of parameters [2].

2

We are unaware of any general dimension count for matrix pencils in the literature. One partial result of Waterhouse [13] counts the codimension of a singular pair of n by n matrices (i.e. the square case) to be $n + 1$.

Our intended application for this work is understanding the occasional failures of existing staircase algorithms to find the "right" Jordan or Kronecker form. The goal of these algorithms is to perturb the input matrix (or pencil) by some small amount so as to give it as much structure as possible, i.e. have as high a codimension as possible. The algorithm is said to **fail** if there is another equally small perturbation which would raise the codimension even further. Therefore, we need to understand how the algorithm produces outputs of each codimension, which is explained in this paper. This is why we need to prove a known result (Theorem 2.1) using a new technique: staircase form. We believe the dimension count for the matrix pencil case (Theorem 2.2) is new.

2   Main Results

Our first result is a new proof of a classical result:

Theorem 2.1 *The codimension of the orbit of a given matrix A is*

$$c_{\text{Jor}} = \sum_{\lambda}(q_1(\lambda) + 3q_2(\lambda) + 5q_3(\lambda) + \ldots),$$

where $q_1(\lambda) \geq q_2(\lambda) \geq q_3(\lambda) \geq \ldots$, denotes the sizes of the Jordan blocks of A corresponding to λ.

Our second result concerns matrix pencils. Before providing the proofs, we review the Jordan and Kronecker structures so that we can fix notation and express this theorem more compactly (Equation (6)) in the next section.

Theorem 2.2 *The codimension of the orbit of $A - \lambda B$ depends only on its Kronecker structure. This codimension can be computed as the sum of separate codimensions as given in Table 1:*

$$c_{\text{Total}} = c_{\text{Jor}} + c_{\text{Right}} + c_{\text{Left}} + c_{\text{Jor,Sing}} + c_{\text{Sing}}$$

3

1. The codimension of the Jordan structure:

$$c_{\text{Jor}} = \sum_{\lambda}(q_1(\lambda) + 3q_2(\lambda) + 5q_3(\lambda) + \dots),$$

where the sum is over all eigenvalues as in Theorem 2.1, including any infinite eigenvalue as well.

2. The codimension of the L singular blocks:

$$c_{\text{Right}} = \sum_{j>k}(j + k - 1),$$

where the sum is taken over all pairs of blocks L_j and L_k for which $j > k$.

3. The codimension of the L^T singular blocks:

$$c_{\text{Left}} = \sum_{j>k}(j + k - 1),$$

where the sum is taken over all pairs of blocks L_j^T and L_k^T for which $j > k$.

4. The codimensions due to interactions of the Jordan structure with the singular blocks:

$$c_{\text{Jor,Sing}} = (\text{size of Jordan structure})(\text{number of singular blocks}).$$

Here the number of singular blocks counts both the left and the right blocks.

5. The codimensions due to interactions between L and L^T singular blocks:

$$c_{\text{Sing}} = \sum_{j,k}(j + k + 2),$$

where the sum is take over all pairs of blocks L_j and L_k^T.

Table 1: Breakdown of the codimension count:

$$c_{\text{Total}} = c_{\text{Jor}} + c_{\text{Right}} + c_{\text{Left}} + c_{\text{Jor,Sing}} + c_{\text{Sing}}$$

4

3 Review of Jordan and Kronecker Canonical Forms and Notation

Some of the basic notation in this area has been reinvented by many authors. So as to make this work self-contained and also to fix notation, we review the basic definitions. Further information may be found in standard matrix theory texts such as [3] or [10].

Given a matrix A that has only one eigenvalue λ it is always possible to find a similarity that transforms A into the form

$$J^\lambda(A) = \mathrm{diag}(J^\lambda_{q_1}, J^\lambda_{q_2}, \ldots) \tag{3}$$

where J^λ_q is a q by q matrix with λ on the diagonal and 1 on the superdiagonal known as a **Jordan block**.

For an arbitrary matrix, it is always possible to find a similarity that transforms A into a union of blocks of the form (3):

$$J(A) = \mathrm{diag}(J^{\lambda_1}(A), J^{\lambda_2}(A), \ldots), \tag{4}$$

where λ_1, λ_2 denotes the distinct eigenvalues of A.

To fix the order of the Jordan blocks within (3), we assume

$$q_1(\lambda) \geq q_2(\lambda) \geq \ldots,$$

but we do not fix the order of the eigenvalues:

Definition 3.1 *The matrix $J(A)$ defined up to eigenvalue orderings is known as the* **Jordan Canonical Form** *of A.*

Definition 3.2 *The sequence of numbers $(q_i(\lambda))$ defined above gives the sizes of the Jordan blocks for the eigenvalue λ. They are known as the* **Segre characteristics** *of A relative to λ.*

It is sometimes convenient to think of this as an infinite sequence with $q_j(\lambda) = 0$ for $j >$ (the number of Jordan blocks corresponding to λ).

Definition 3.3 *The* **elementary divisors** *of the matrix $A - xI$ are the polynomials $(\lambda - x)^{q_i(\lambda)}$ in the indeterminate x, where λ is an eigenvalue of A and $q_i(\lambda)$ is a Segre characteristic corresponding to λ.*

Definition 3.4 *The* **invariant factors** *of the matrix* $A - xI$ *are the polynomials* $P_i(x) = \prod_\lambda (\lambda - x)^{q_i(\lambda)}$. *It follows that if we let* p_i *denote the degree of the ith invariant factor then*

$$p_i = \sum_\lambda q_i(\lambda).$$

Of course $n = \sum p_i$ because this counts the sizes of all the Jordan blocks of all the eigenvalues of A.

Some authors (see [10] pages 43 and 93) consider the quantity m_i defined as the degree of the greatest common divisor of all the i by i minors of the linear matrix polynomial $A - \lambda I$. It can be shown that $m_i = p_{n+1-i} + \ldots + p_n$.

Definition 3.5 *The* **nullity** *of an n by n matrix A is $n - \text{rank}(A)$. For m by n matrices the row nullity and the column nullity are $m - \text{rank}(A)$ and $n - \text{rank}(A)$ respectively.*

Definition 3.6 *Let $w_j(\lambda)$ denote the difference*

$$\text{nullity}(A - \lambda I)^j - \text{nullity}(A - \lambda I)^{j-1}.$$

The numbers $w_k(\lambda)$ are the **Weyr characteristics** *of A relative to λ. The number of blocks $J_q(\lambda)$ with $q \geq j$ is exactly $w_j(\lambda)$.*

Let $A - \lambda B$ be an m by n matrix pencil. It is possible to find an equivalent pencil $K(A - \lambda B)$ in the Kronecker Form:

$$K(A - \lambda B) = \text{diag}(L_{\epsilon_1}, \ldots, L_{\epsilon_g}, L_{\eta_1}^T, \ldots, L_{\eta_h}^T, J, J^\infty). \tag{5}$$

The L_ϵ blocks are ϵ by $\epsilon + 1$ rectangular blocks with λ on the diagonal and 1 on the superdiagonal. The L_η^T blocks are $\eta + 1$ by η, with λ on the diagonal, and 1 on the subdiagonal. The ϵ and η can be 0, leading to 0 columns and rows respectively. The J block is of the form (4) with the addition of λI. This constitutes the Jordan structure of the finite eigenvalues. The J^∞ block is the union of blocks of size $q_i(\infty)$ each of which has 1 on the main diagonal and λ on the superdiagonal. This constitutes the Jordan structure corresponding to the infinite eigenvalue. Frequently there will be no need to distinguish between the finite and infinite eigenvalues.

The L and L^T blocks constitute the **singular** part of the pencil. The Jordan structure for finite and infinite eigenvalues constitutes the **regular** part of the pencil. The Segre characteristics remain well defined for a matrix pencil, but we must include the characteristics for the infinite eigenvalue as well.

6

Definition 3.7 *Let*

$$0 \le \epsilon_1 \le \epsilon_2 \le \ldots \le \epsilon_g$$

denote the sizes of the g L blocks of a pencil, and let

$$0 \le \eta_1 \le \eta_2 \le \ldots \le \eta_h$$

denote the sizes of the h L^T blocks. Then the numbers ϵ_i are known as the **column minimal indices**, *while the η_i are the* **row minimal indices**.

We can now recast Theorem 2.2 using the notation from the previous definitions. The codimension of the orbit of $A - \lambda B$ can be written compactly as

$$\begin{aligned}
\mathrm{cod}(\mathrm{orbit}(A - \lambda B)) = {}& (p_1 + 3p_2 + 5p_3 + \ldots) + (g + h)\sum p_i \\
& + \sum_{i>j}(\epsilon_i + \epsilon_j - 1) + \sum_{i>j}(\eta_i + \eta_j - 1) + \sum_{i,j}(\epsilon_i + \eta_j + 2), \qquad (6)
\end{aligned}$$

where the p_i include any infinite eigenvalue blocks.

4 Proofs of Theorem 2.1

4.1 Classical Proof

Consider conjugating the matrix A by $I + \delta X$, where δ is a small real number. This yields

$$A + \delta(XA - AX) + O(\delta^2),$$

from which it is evident that the tangent space to orbit(A) at A consists of the matrices of the form $XA - AX$. The dimension of the orbit is equal to the dimension of the tangent space so that the codimension of the orbit is equal to the dimension of the nullspace of the mapping that sends X to $XA - AX$. The codimension of the orbit is then the number of linearly independent solutions to $AX = XA$. This number of solutions is well known to be

$$p_1 + 3p_2 + 5p_3 + \ldots.$$

(See page 222 of volume 1 of [3].) $\qquad\qquad\Box$

An alternative expression for the number of solutions to $AX = XA$ is

$$n + 2(m_1 + \ldots + m_{n-1})$$

as given in [10]. According to the remark following Definition 3.4, these expressions are identical.

4.2 Proof Based on the Staircase Algorithm

The staircase algorithm for the computation of the Jordan Canonical Form appears in [4, 5, 8, 9, 11].

Let A be a matrix and let λ be an eigenvalue of A. The staircase algorithm computes a matrix that is orthogonally similar to $A - \lambda I$, but with the staircase form as illustrated with the following example:

	w_1	w_2	w_3	w_4	n'
w_1	0	A_{12}	*	*	*
w_2		0	A_{23}	*	*
w_3			0	A_{34}	*
w_4				0	A_{45}
n'					A'

In the above example, the superdiagonal blocks $A_{i,i+1}$ are of full rank, while the staircase region in the lower triangle is entirely 0. If A has only one eigenvalue λ then n' is 0 and the last block row and block column do not appear. If A has other eigenvalues, then the staircase form corresponding to the remaining eigenvalues may be extracted from A'.

An important fact [10, p. 74] is that

Lemma 4.1 *The w_i computed by the staircase algorithm for the eigenvalue λ are the Weyr characteristics corresponding to the eigenvalue λ.*

Definition 4.1 *Let $k_1 \geq k_2 \geq k_3 \geq \ldots \geq 0$ be a partition of the positive integer k (i.e. $k = k_1 + k_2 + \ldots$). Let l_j denote the number of k_i that are greater than or equal to j. Then the l_j form a partition of k known as the* **conjugate partition** *of the k_i.*

It is easy to verify that the property of being a conjugate partition is symmetric. For example, $17=6+6+3+1+1=5+3+3+2+2+2$ are conjugate partitions of 17. This is easy to verify by reading the diagram below (known as a Ferrers diagram) vertically and horizontally:

	6	6	3	1	1
5	.	.	.	.	.
3	.	.	.		
3	.	.	.		
2	.	.			
2	.	.			
2	.	.			

The idea of the conjugate partition is very simple, yet very powerful. It allows the interchange of summations:

$$\sum_i \sum_j^{k_i} f(i,j) = \sum_j \sum_i^{l_j} f(i,j),$$

where $f(i,j)$ is any function of i and j, and the k_i and l_j are conjugate partitions. In particular, the Weyr characteristics and the Segre characteristics of a matrix corresponding to a particular eigenvalue are conjugate partitions.

We will need one more lemma concerning the dimension of a rectangular matrix with a particular rank:

Lemma 4.2 *The codimension of the set of m by n matrices with rank r is $(m-r)(n-r)$, the product of the row and column nullities. When $m = n$ the codimension is then the square of the nullity.*

Proof There are $(m-r)(n-r)$ numbers determined by the condition that every $r+1$ rowed minor is 0. $\qquad\square$

Proof of Theorem 2.1:

We remark that the proof of this theorem is deceptively simple because of the nice algebraic machinery available.

The ith step of the staircase algorithm computes the Weyr characteristic w_i for a given eigenvalue λ by computing the nullity of the lower right submatrix of a transformed A obtained by deleting the first $w_1 + \ldots + w_{i-1}$ rows and columns of A. This further restricts the Jordan form of the matrix to a set of higher codimension. The incremental increase in codimension is given by Lemma 4.2 as w_i^2. The total codimension due to the eigenvalue λ is then

$$\sum_i w_i^2 \;=\; \sum_i \sum_{k=1}^{w_i} (2k - 1)$$

9

$$= \sum_k \sum_{i=1}^{q_k} (2k - 1)$$

$$= \sum_k (2k - 1) q_k,$$

using the fact that the Weyr and Segre characteristics are conjugate partitions. The total codimension is obtained by summing over all the eigenvalues. $\square$

5 Proof of Theorem 2.2

We include two proofs both of which we believe to be new. The first proof requires counting the number of independent solutions to two simultaneous matrix equations derived by analyzing the tangent space, while the second proof requires an analysis of the staircase algorithms for the Kronecker canonical form.

5.1 Proof based on the tangent space

Consider an orbit preserving transformation of the m by n pencil $A - \lambda B$ obtained by multiplying on the left by $I + \delta X$ and the right by $I - \delta Y$, where δ is a small real number. This yields $A - \lambda B + \delta(X(A - \lambda B) - (A - \lambda B)Y) + O(\delta^2)$, from which it is evident that the tangent space to the orbit of the pencil consists of the pencils that can be represented in the form

$$f(X, Y) = X(A - \lambda B) - (A - \lambda B)Y, \tag{7}$$

where X is an m by m matrix and Y is an n by n matrix.

Since (7) maps a space of dimension $m^2 + n^2$ linearly into a space of dimension $2mn$, the dimension of the image space is $m^2 + n^2 - d$, where d is the dimension of the kernel of $f(X, Y)$, and so the codimension is

$$2mn - (m^2 + n^2 - d) = d - (m - n)^2. \tag{8}$$

As in the Jordan case, we need to calculate d, the number of linearly independent solutions to $f(X, Y) = 0$. This can be written as the two simultaneous equations

$$XA = AY \text{ and } XB = BY. \tag{9}$$

10

Unfortunately, we can not simply quote a classical count of the number of independent solutions to (9) as we were able to do in Section 4.1. However since

$$Pf(X,Y)Q^{-1} = (PXP^{-1})P(A - \lambda B)Q^{-1} - P(A - \lambda B)Q^{-1}(QYQ^{-1}),$$

it follows that the number of linearly independent solutions to $f(X,Y) = 0$ depends only on the Kronecker structure of $A - \lambda B$. Thus, we assume that $A - \lambda B$ is already in Kronecker canonical form $M = \text{diag}(M_1, M_2, \ldots)$. The Kronecker case is more complicated than the Jordan case due to the greater number of possibilities for the Kronecker structure M.

We partition the equation $XM = MY$ conformally with $M = \text{diag}(M_1, M_2, \ldots)$ so that $X_{ij}M_j = M_i Y_{ij}$, where M_k is m_k by n_k X_{ij} is m_j by m_i and Y_{ij} is n_j by n_i. The next lemma allows us to compute the quantity d in Equation (9) as the sum of the number d_{ij} of independent solutions of $X_{ij}M_j = M_i Y_{ij}$ in the variables X_{ij} and Y_{ij}.

Lemma 5.1 *In terms of the above notation*

$$d = \sum_{i,j} d_{ij}.$$

Proof The proof is evident from the example

$$\begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} \begin{pmatrix} M_1 & \\ & M_2 \end{pmatrix} = \begin{pmatrix} M_1 & \\ & M_2 \end{pmatrix} \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix}.$$

$\square$

Given any two blocks, M_i and M_j (we allow $i = j$ here) we define their **interaction** and the **cointeraction**:

Definition 5.1 *Let M_i be $m_i \times n_i$ and let M_j be $m_j \times n_j$. Let X be an arbitrary $m_j \times m_i$ matrix and Y be an arbitrary $n_j \times n_i$ matrix. We define the **interaction** d_{ij} of M_i with M_j as the dimension of the linear space X,Y such that $XM_i = M_jY$. We define the **cointeraction** of M_i with M_j as $d_{ij} + 2m_i n_j - (m_i m_j + n_i n_j)$.*

This definition makes sense if we consider

Lemma 5.2 *The codimension of a matrix pencil M with Kronecker structure $\text{diag}(M_1, M_2, \ldots)$ is the sum of cointeractions of M_i with M_j for all combinations of i and j.*

Proof The sum of the cointeractions is

$$\sum_{i,j} \{d_{ij} + 2m_i n_j - (m_i m_j + n_i n_j)\}$$

$$= \sum_{ij} d_{ij} + 2\sum_i m_i \sum_j n_j - \left((\sum_i m_i)^2 + (\sum_j n_j)^2 \right)$$

$$= d + 2mn - (m^2 + n^2)$$

as in Equation (8). □

We must now count the number of linearly independent solutions to the following equations:

- $XL_j = L_k Y$ and $XL_j^T = L_k^T Y$

- $XL_j = L_k^T Y$ and $XL_j^T = L_k Y$

- $XL_j = JY$

- $XJ = JY$

where J denotes the non-singular structure of the pencil. From this information, we compute the cointeractions.

5.1.1 $XL_j = L_k Y$ and $XL_j^T = L_k^T Y$

Consider the equation $XL_j = L_k Y$, where X is an unknown k by j matrix and Y is an unknown $k+1$ by $j+1$ matrix. This equation is equivalent to the two equations

$$X[0 \; I_j] = [I_k \; 0]Y$$
$$X[I_j \; 0] = [0 \; I_k]Y,$$

where 0 denotes a column of zeros. These two equations are in turn equivalent to the conditions

$$Y_{\alpha,\beta} = Y_{\alpha+1,\beta+1}, \; \alpha = 1,\ldots,k, \; \beta = 1,\ldots,j$$
$$Y_{\alpha+1,1} = Y_{\alpha,j+1} = 0, \; \alpha = 1,\ldots,k$$
$$X_{\alpha,\beta} = Y_{\alpha,\beta+1}, \; \alpha = 1,\ldots,k, \; \beta = 1,\ldots,j$$

12

If $k < j$ there is only the trivial solution $X = 0$ and $Y = 0$. If $j \geq k$ then there are non-trivial solutions: Y can be any upper triangular Toeplitz matrix with $1 + j - k$ diagonals starting from the main diagonal. X is then obtained from Y by ommiting the first row and column. The interaction of L_j with L_k is $1 + j - k$.

When $j \geq k$ the cointeraction of L_j with L_k is

$$2(j + 1)k - (jk + (k + 1)(j + 1)) + (1 + j - k) = 0,$$

while if $k > j$ the cointeraction is

$$2(j + 1)k - (jk + (k + 1)(j + 1)) = k - j - 1.$$

By symmetry, we obtain the same result for blocks of the form L_j^T. We also remark that this analysis is correct even if j or k is 0.

5.1.2 $XL_j = L_k^T Y$ and $XL_j^T = L_k Y$

The equation $XL_j = L_k^T Y$ has only the trivial solution $X = Y = 0$ so that the interaction of L_j with L_k^T is 0. The cointeraction is then

$$2(j + 1)(k + 1) - (j(k + 1) + (j + 1)k) = j + k + 2.$$

Using similar techniques it is possible to show that the cointeraction of L_j^T with L_k is 0.

5.1.3 Jordan Blocks and Singular Blocks

Let J_k be a single Jordan block of size k corresponding to a finite or infinite eigenvalue. We omit the tedious algebra, which is analogous to that in subsection 5.1.1, and state the conclusion that the cointeraction of J_k with L_j is k while the reverse cointeraction is 0.

5.1.4 Jordan Blocks with other Jordan Blocks

The only difference between the non-singular portion of the Kronecker structure and the Jordan structure of a single matrix is the possibility of an infinite eigenvalue. Again we omit the tedious algebra, but it is possible to show that an infinite eigenvalue behaves exactly as if it were finite. Thus the cointeractions of the non-singular portion of the pencil with itself is exactly as in Theorem 2.1.

5.1.5 Proof of Theorem 2.2

The proof follows from the analysis of the cases presented in Sections 5.1.1 through 5.1.4.   □

5.2 Proof Based on the Staircase Algorithm

We begin by reviewing the staircase algorithm. The version we use has three passes. Let $A - \lambda B$ be an m by n matrix pencil. The first pass produces two sequences of numbers s_i and r_i and returns a pencil $A' - \lambda B'$ with no L_j blocks and no zero eigenvalues. The sequence satisfies

$$s_0 \geq r_0 \geq s_1 \geq r_1 \geq s_2 \geq \ldots,$$

where

- $s_i - r_i =$ the number of L_i blocks and

- $r_i - s_{i+1} =$ the number of J_{i+1}^0 blocks.

The algorithm is as follows:

1. Compute $s_0 =$ the column nullity of A.

2. Postmultiply the pencil by an orthogonal Q so that the first s_0 columns of A vanish.

3. Compute $r_0 =$ the rank of the first s_0 columns of the transformed B

4. Premultiply the transformed pencil by an orthogonal P so that the lower left $m - r_0$ by s_0 submatrix of the transformed B is zero.

5. Repeat this on the subpencil in the lower right $m_1 = m - r_0$ rows and $n_1 = n - s_0$ columns of the transformed pencil, defining s_1, r_1, m_2, n_2, etc.

We illustrate with the following small example:

	s_0	s_1	s_2	s_3	n'
r_0	$0 - \lambda B_{00}$	$A_{01} - \lambda B_{01}$	*	*	*
r_1		$0 - \lambda B_{11}$	$A_{12} - \lambda B_{12}$	*	*
r_2			$0 - \lambda B_{22}$	$A_{23} - \lambda B_{23}$	*
m'					$A' - \lambda B'$

14

On completion, the B_{ii} blocks have full row rank, and the $A_{i,i+1}$ blocks have full column rank.

The ith step of this algorithm (starting from $i = 0$) works on the m_i by n_i lower right subpencil of $A - \lambda B$ and computes the indices r_i and s_i. Just as in the single matrix case, each step restricts the Kronecker form of the pencil to a set of higher codimension. The incremental increase in codimension is given by Lemma 4.2, as the sum of the products of the row and column nullities of submatrices of A and B. Specifically the m_i by n_i submatrix of A has column nullity s_i, rank $n_i - s_i$, row nullity $m_i + s_i - n_i$, and so by Lemma 4.2 codimension $(m_i + s_i - n_i)s_i$. Similarly the codimension due to B at step i is $(m_i - r_i)(s_i - r_i)$. The first pass through the algorithm determines the L and J^0 blocks so that the codimension due to these blocks is given by

$$\sum_i \left\{ (m_i + s_i - n_i)s_i + (m_i - r_i)(s_i - r_i) \right\}. \tag{10}$$

We proceed to show that (10) is the formula given in Theorem 2.2.

For convenience we list our notation:

m_i number of rows in the lower right subpencil at step $i = m - \sum_{k=1}^{i-1} r_k$

n_i number of columns in the lower right subpencil at step $i = n - \sum_{k=1}^{i-1} s_k$

s_i column nullity of A-part of subpencil at step i

r_i row rank of the first s_i columns of B-part of subpencil at step i

l_i number of L_i blocks in the original pencil

l_i' number of L_i^T blocks in the original pencil

t_i number of J_i^0 blocks in the original pencil

u size of the regular structure corresponding to $\lambda \neq 0$.

5.2.1   Only left singular blocks

We begin by assuming that our pencil only contains left singular blocks. Let l_i denote the number of L_i blocks. It is easy to show by induction that the algorithm computes

$$m_i = \sum_{j=i+1}^{\infty} (j - i)l_j$$

$$n_i = \sum_{j=i+1}^{\infty} (1 + j - i)l_j$$

15

$$s_i = \sum_{j=i}^{\infty} l_j$$

$$r_i = \sum_{j=i+1}^{\infty} l_j.$$

Thus for this case expression (10) evaluates to

$$\alpha = \sum_{i=0}^{\infty} l_i \sum_{j=i+1}^{\infty} (j - i - 1)l_j. \tag{11}$$

This is exactly $\sum_{i>j}(\epsilon_i + \epsilon_j - 1)$ as in (6).

5.2.2 Left singular blocks and J^0 blocks

We now add the assumption that there are J^0 blocks as well. Let t_i be the number of J_i^0 blocks, i.e., Jordan blocks of size i corresponding to a zero eigenvalue. Again by induction it is possible to show

$$m_i = \sum_{j=i+1}^{\infty} (j - i)(l_j + t_j)$$

$$n_i = m_i + \sum_{j=i}^{\infty} l_j$$

$$s_i = \sum_{j=i}^{\infty} l_j + \sum_{j=i+1}^{\infty} t_j$$

$$r_i = \sum_{j=i+1}^{\infty} (l_j + t_j).$$

Now for this case Expression (10) evaluates to

$$\beta = \sum_{i=0}^{\infty} \left\{ \left(\sum_{j=i+1}^{\infty} t_j \right) \left(\sum_{j=i}^{\infty} l_j + \sum_{j=i+1}^{\infty} t_j \right) + l_i \sum_{j=i+1}^{\infty} (j - i - 1)(l_j + t_j) \right\} \tag{12}$$

which can readily be manipulated to be

$$\beta = \alpha + \sum_{i=0}^{\infty} \left(\sum_{j=i+1}^{\infty} t_j \right)^2 + \sum_{i=0}^{\infty} \left\{ \sum_{j=i}^{\infty} l_j \sum_{k=i+1}^{\infty} t_k + l_i \sum_{j=i+1}^{\infty} (j - i - 1)(l_j + t_j) \right\},$$

16

where α is the same interaction among the left singular blocks as in Equation (11). We recognize $(\sum_{j=i+1}^{\infty} t_j)^2$ as the square of the $i+1$'st Weyr characteristic of the infinite eigenvalue. From our new proof of Theorem 2.1 we know that this sum is the codimension due to the infinite eigenvalue alone.

Lastly, we must evaluate

$$\sum_{i=0}^{\infty}\left\{\sum_{j=i}^{\infty}l_j\sum_{k=i+1}^{\infty}t_k + l_i\sum_{j=i+1}^{\infty}(j-i-1)(l_j+t_j)\right\}$$

$$=\sum_{i=0}^{\infty}l_i(\sum_{j=0}^{i}\sum_{k=j+1}^{\infty}t_k + \sum_{k=i+1}^{\infty}(k-i-1)t_k)$$

$$=\sum_{i=0}^{\infty}l_i(\sum_{k=1}^{i}\sum_{j=0}^{k-1}t_k + \sum_{k=i+1}^{\infty}\sum_{j=0}^{i}t_k + \sum_{k=i+1}^{\infty}(k-i-1)t_k)$$

$$=(\sum_{i=0}^{\infty}l_i)(\sum_{k=1}^{\infty}kt_k)$$

$$= (\text{size of Jordan structure for } \lambda = 0)(\text{number of left singular blocks})$$

5.2.3 Arbitrary singular blocks and arbitrary Jordan structure

We complete the first pass through the algorithm by defining l_i' to denote the number of L_i^T blocks, and u to be the size of the regular Jordan structure for $\lambda \neq 0$. Thus, u includes the structre for $\lambda = \infty$ which plays no special role during the first pass through the algorithm.

We once again omit the details, but it is possible to show by induction that the algorithm computes

$$m_i = m_i^0 + \sum_{j=0}^{\infty}(j+1)l_j' + u$$

$$n_i = n_i^0 + \sum_{j=0}^{\infty}jl_j' + u$$

$$s_i = s_i^0$$

$$r_i = r_i^0,$$

where the superscript 0 indicates no right singular structure and no non-zero regular structure, i.e. as in the notation of Section 5.2.

We now have that the codimension expression in (10) is

$$\gamma = \beta + \sum_{i=0}^{\infty} \left\{ (\sum_{j=0}^{\infty} l'_j)(\sum_{j=i}^{\infty} l_j + \sum_{j=i+1}^{\infty} t_j) + l_i(\sum_{j=0}^{\infty}(j+1)l'_j + u) \right\},$$

where β is as in (12). With some algebraic manipulation, we obtain

$$\gamma = \beta + \sum_{i,j=0}^{\infty} l_i l'_j(i+j+2) + u \sum_{i=0}^{\infty} l_j + (\sum_{i=0}^{\infty} l'_i)(\sum_{k=1}^{\infty} k t_k).$$

The terms here are the terms

$$\sum_{i,j}(\epsilon_i + \eta_j + 2) + g \sum_i (p_i - q_i^0) + h \sum_i q_i^0.$$

5.3 Second and third passes through algorithm

The first pass through the algorithm gives us a pencil $A' - \lambda B'$, which may have only L_j^T blocks and nonzero eigenvalues. We then run the algorithm on $(B' - \lambda A')^T$, so that the indices that gave the right singular blocks before now give the left singular blocks. The indices that described $\lambda = 0$ now describe $\lambda = \infty$. This algorithm returns a pencil with only a regular part that has no zero or infinite eigenvalues.

If we reinvoke the previous results, we see that the second pass through the algorithms nearly completes the entire expression (6). The only gap is

$$\sum_{\lambda \notin \{0,\infty\}} (q_1^{\lambda} + 3q_2^{\lambda} + 5q_3^{\lambda} + \ldots).$$

This is just the Jordan structure of the regular part other than the zero and infinite eigenvalues. This is covered in the third phase of the algorithm, completing the proof. □

6 Observations About Genericity and the Waterhouse Theorems

As corollaries to Theorem 2.2, we comment on the most generic structure for an m by n pencil, as well as the most generic structure for an n by n singular pencil.

18

Corollary 6.1 *The generic Kronecker structure for a matrix pencil with* $d = n - m > 0$ *has the form*

$$diag(L_\alpha, \ldots, L_\alpha, L_{\alpha+1}, \ldots, L_{\alpha+1}),$$

where $\alpha = \lfloor m/d \rfloor$, *the total number of blocks is* d, *while the number of* $L_{\alpha+1}$ *blocks is given by* $m \bmod d$ *(which is 0 when d divides m).*

Proof From Theorem 2.2, this structure (and only this structure) has codimension 0, i.e., it is generic. □

The same statement holds when $d = m - n > 0$ if we replace the L_α and $L_{\alpha+1}$ blocks by their transposes.

Corollary 6.1 was obtained by Van Dooren, Wilkinson, and Wonham as discussed on page 3.55 of [12].

From Theorem 2.2 square pencils are generically non-singular. As was observed in [13], the smallest codimension for singular pairs is $n + 1$. (A simple argument would be that there are $n + 1$ conditions that the the $n + 1$ coefficients of λ in $\det(A - \lambda B) = 0$.) Another corollary to our theorem reproduces a result of Waterhouse([13]:

Corollary 6.2 *The generic singular pencils of size n by n have Kronecker structures*

$$diag(L_j, L^T_{n-j-1}),$$

where $j = 0, \ldots, n - 1$.

Proof Only these pencils have a singular part and also have codimension $n + 1$. □

More generally, [13] has shown that if a square matrix has one L_r block and one L^T_s block and otherwise has a generic $n - r - s - 1 \times n - r - s - 1$ block (eigenvalues unspecified), then the codimension is $(r + s + 2) + 2(n - r - s - 1) = 2n - (r + s)$. This too readily follows from our results.

References

[1] V.I. Arnold, On matrices depending on parameters, *Russian Math. Surveys 26* (1971), 29–43.

[2] T. Fairgrieve, *The application of singularity theory to the computation of the Jordan canonical form*, unpublished M.Sc. thesis, Computer Science Dept., University of Toronto, Canada (1986)

[3] F.R. Gantmacher, *The Theory of Matrices*, Vols. 1, 2, Chelsea, New York, 1959.

[4] G.H. Golub and C.F. van Loan, *Matrix Computations, Second Edition*, Johns Hopkins University Press, Baltimore, 1989.

[5] G.H. Golub and J.H. Wilkinson, Ill-conditioned eigensystems and the computation of the Jordan canonical form, *SIAM Review 18* (1976), 578–619

[6] B. Kågström, The generalized singular value decomposition and the general $A - \lambda B$ problem, *BIT* 24 (1985), 568–583.

[7] B. Kågström, RGSVD: an algorithm for computing the Kronecker structure of reducing subspaces of singular $A - \lambda B$ pencils, *SIAM J. Sci. Stat. Comp.* 7 (1986), 185–211.

[8] B. Kågström and A. Ruhe, An algorithm for numerical computation of the Jordan normal form of a complex matrix, *ACM Trans. Math. Soft. 6* (1980), 398–419.

[9] V.N. Kublanovskaya, On a method of solving the complete eigenvalue problem for a degenerate matrix, *USSR Comp. Math. & Math. Phys. 6* (1968), 1–14.

[10] C.C. MacDuffee, *The Theory of Matrices*, Springer-Verlag, Berlin, 1933.

[11] A. Ruhe, An algorithm for the numerical determination of the structure of a general matrix, *BIT 10,* (1970), 196–216.

[12] P. Van Dooren, The computation of Kronecker's canonical form of a singular pencil, *Lin. Alg. Appl. 27* (1979), 103-40

[13] W.C. Waterhouse, The codimension of singular matrix pairs, *Lin. Alg. Appl. 57* (1984), 227–245.

[14] J. H. Wilkinson. Linear differential equations and Kronecker's canonical form, in *Recent Advances in Numerical Analysis*, eds. C. de Boor and G. Golub, Academic Press, New York, (1978), 231–265.

[15] J. H. Wilkinson. Kronecker's canonical form and the QZ algorithm, *Lin. Alg. Appl. 28* (1979), 285–303.

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
TECHNICAL INFORMATION DEPARTMENT
BERKELEY, CALIFORNIA 94720