SIAM J. MATRIX ANAL. APPL.
Vol. 20, No. 3, pp. 667–699
© 1999 Society for Industrial and Applied Mathematics

A GEOMETRIC APPROACH TO PERTURBATION THEORY OF MATRICES AND MATRIX PENCILS. PART II: A STRATIFICATION-ENHANCED STAIRCASE ALGORITHM*

ALAN EDELMAN†, ERIK ELMROTH‡, AND BO KÅGSTRÖM‡

Abstract. Computing the Jordan form of a matrix or the Kronecker structure of a pencil is a well-known ill-posed problem. We propose that knowledge of the closure relations, i.e., the stratification, of the orbits and bundles of the various forms may be applied in the staircase algorithm. Here we discuss and complete the mathematical theory of these relationships and show how they may be applied to the staircase algorithm. This paper is a continuation of our Part I paper on versal deformations, but it may also be read independently.

Key words. Jordan canonical form, Kronecker canonical form, staircase algorithm, matrix pencils, closure relations, stratification, quivers

AMS subject classifications. 65F15, 15A21, 15A22

PII. S0895479896310184

Notation.

A	A square matrix of size $n \times n$. I or I_n is the identity matrix.
A^T	The transpose of A.
$\mathrm{vec}(A)$	An ordered stack of the columns of a matrix A from left to right.
$\det(A)$	Determinant of A.
$\mathcal{N}(A)$	The column nullspace of A.
$\mathrm{diag}(A_1, \ldots, A_b)$	A block diagonal matrix with diagonal blocks A_i.
$A \otimes B$	The Kronecker product of two matrices A and B whose (i,j)th block element is $a_{ij}B$.
$A_1 \boxplus A_2$	A canonical form whose Segre characteristics are the sum of those of A_1 and A_2, or equivalently whose Weyr characteristics are the union of those of A_1 and A_2.
$A - \lambda B$	A matrix pencil of size $m \times n$. Also denoted P.
$A^{(k)} - \lambda B^{(k)}$	Deflated pencil at step k of a staircase algorithm.
m_k, s_k	$m_k =$ dimension of nullspace of $A^{(k)}$, $m_k - s_k =$ dimension of common nullspace of $A^{(k)}$ and $B^{(k)}$.
λ_i	Eigenvalue of A or $A - \lambda B$. Also μ_i, α, β, γ, and δ are used.
$J_j(\lambda_i)$	Jordan block of size $j \times j$ associated with λ_i.
N_j	Jordan block of size $j \times j$ associated with the infinite eigenvalue (sometimes denoted $J_j(\infty)$).
L_j	Singular block of right (column) minimal index of size $j \times (j+1)$.
L_j^T	Singular block of left (row) minimal index of size $(j+1) \times j$.

---

*Received by the editors October 1, 1996; accepted for publication (in revised form) by J. Varah May 17, 1998; published electronically March 23, 1999.
http://www.siam.org/journals/simax/20-3/31018.html

†Department of Mathematics, Room 2-380, Massachusetts Institute of Technology, Cambridge, MA 02139 (edelman@math.mit.edu). The work of this author was supported by NSF grant DMS-9501278 and an Alfred P. Sloan Foundation Research Fellowship.

‡Department of Computing Science, Umeå University, S-901 87 Umeå, Sweden (elmroth@cs.umu.se, bokg@cs.umu.se). The work of the authors was supported by Swedish Research Council for Engineering Sciences grant TFR 222-95-34.

$\alpha^i\alpha^j\beta^k$	Compact notation for $J_i(\alpha) \oplus J_j(\alpha) \oplus J_k(\beta)$.
$\mathrm{nrk}(A - \lambda B)$	Normal rank of $A - \lambda B$.
$\mathcal{O}(A)$	The orbit of A, i.e., the set of matrices similar to A.
$\mathcal{O}(A - \lambda B)$	The orbit of $A - \lambda B$, i.e., the set of matrix pencils equivalent to $A - \lambda B$.
$\overline{\mathcal{O}}(\cdot)$	The closure of an orbit.
$\mathcal{B}(A)$	The bundle of A, i.e., the set of matrices with the Jordan structure of A, but with the eigenvalues unspecified.
$\mathcal{B}(A - \lambda B)$	The bundle of $A - \lambda B$, i.e., the set of matrix pencils with the Kronecker structure of $A - \lambda B$, but with the eigenvalues unspecified.
$\overline{\mathcal{B}}(\cdot)$	The closure of a bundle.
$\mathcal{S} \oplus \mathcal{T}$	Direct sum of subspaces $\mathcal{S}$ and $\mathcal{T}$ of $\mathbf{R}^n$.
$\dim(\mathcal{S})$	Dimension of subspace $\mathcal{S}$. $\dim(S)$ denotes dimension of subspace spanned by the columns of S.
$\mathrm{cod}(\mathcal{S})$	Codimension is the dimension of the subspace complementary to $\mathcal{S}$.
$\kappa(A)$	$\kappa = (k_1, k_2, \ldots)$ is the integer partition representing the Segre characteristics for an eigenvalue λ_i of A.
$\mu(A)$	Integer partition representing the Weyr characteristics for an eigenvalue λ_i of A.
κ'	The conjugate partition of κ, e.g., $\mu(A) = \kappa(A)'$.
$\mathcal{J}_{\mu_i}$	$\mathcal{J}_{\mu_i} = (j_1, j_2, \ldots)$ is the integer partition representing the Weyr characteristics of $A - \lambda B$ for the eigenvalue μ_i.
$\mathcal{R}$	$\mathcal{R} = (r_0, r_1, \ldots)$ is the integer partition representing the right singular structure of $A - \lambda B$.
$\mathcal{L}$	$\mathcal{L} = (l_0, l_1, \ldots)$ is the integer partition representing the left singular structure of $A - \lambda B$.
$\langle P_1, P_2 \rangle_1$	Inner product for Kronecker structures defined as $\dim\{V : A_2VB_1^T = B_2VA_1^T\}$, where $P_1 = A_1 - \lambda B_1$ and $P_2 = A_2 - \lambda B_2$.
$\langle P_1, P_2 \rangle_2$	Inner product for Kronecker structures defined as $\dim\{(U, V) : UP_1 = P_2V\}$ (also denoted $\langle P_1, P_2 \rangle$).
$\mathbf{A}_m$, $\tilde{\mathbf{A}}_m$	The quivers $\mathbf{A}_m$ and $\tilde{\mathbf{A}}_m$.
$(x - y)_+$	$max(x - y, 0)$.

1. Introduction.

The determination of the Jordan form of a matrix A with multiple defective or derogatory eigenvalues is an ill-posed problem in the presence of roundoff error [26]. The same is true for the Kronecker form of a matrix pencil $A - \lambda B$. Therefore modern numerical software such as GUPTRI [15, 16] regularizes the problem by allowing a tolerance for rank decisions so as to find a matrix pencil near $A - \lambda B$ with an interesting Kronecker (or Jordan) structure. These algorithms are known to occasionally fail, thereby accidentally producing wrong nearby structures. Failure appears to occur when the matrix or pencil is close to a manifold of interesting structures of higher codimension [13]. Motivating examples arise in control theory, where linear control systems have been found that the staircase algorithm can decide are easily controllable, but in fact these systems were nearly uncontrollable (see, e.g., [6]). Because of these occasional failures, we propose to make use of the mathematical knowledge of the *stratification* of the Jordan and Kronecker structures in order to enhance the staircase algorithm. This stratification, in effect, shows which structures

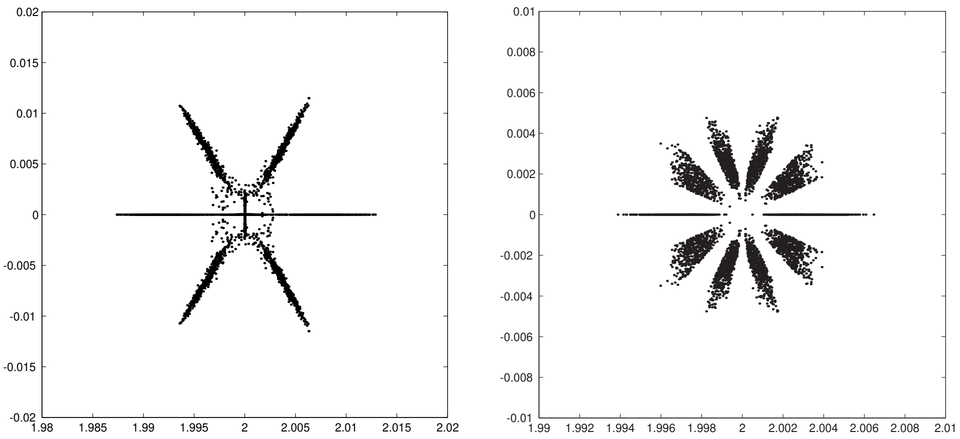


FIG. 1.1. *Perturbed eigenvalues of* $J_3(2) \oplus J_2(2)$. *Left: predominate structure. Right: nearby structure* $J_5(2)$.

are near other structures (in the sense of being in the closure) in the space of matrices. During the execution of a staircase algorithm the user or the program can be aware of the other nearby structures.

There are a number of ways to see the effects of nearby structures. Numerical experiments on random perturbations of 2×3 matrix pencils using GUPTRI are reported by Elmroth and Kågström [21, Table 3.1]. Assuming a fixed relative accuracy of the input data, the structures are studied as functions of the sizes of the perturbations. Even in the admittedly small 2×3 case, it becomes clear that there is an interesting combinatorial relationship among the possibilities, which we will investigate.

Consider the qualitative approach to the Jordan form proposed by Chaitin-Chatelin and Frayssé [10] using their example of the Ruhe matrix whose Jordan structure for eigenvalue $\lambda = 2$ is $J_3(2) \oplus J_2(2)$. The computed eigenvalues of roughly 1000 (real) normally distributed random perturbations of size $\sqrt{\epsilon}$ ($\epsilon = 2^{-52}$ is the usual IEEE double precision "epsilon") allow us to plot perturbed eigenvalues as in the picture to the left in Figure 1.1.

The six lines from the origin (four are fuzzy) and the smaller cross predict the Jordan structure $J_3(2) \oplus J_2(2)$. Besides this predominant structure, other structures may also appear [10]. We ran $50,000$ tests and filtered out those with roughly the predominant structure, thereby leaving around 1000 matrices whose perturbed eigenvalues appear to the right in Figure 1.1. This figure suggests that the structure $J_5(2)$ is somehow nearby. It turns out that $J_4(2) \oplus J_1(2)$ is also nearby but is much rarer. In 500,000 random tests none were found.

In addition to $J_4(2) \oplus J_1(2)$ and $J_5(2)$, one may wonder if we may have somehow missed other nearby structures. (We have not!) A more important question is if an algorithm such as GUPTRI or the qualitative approach suggest a certain Jordan structure, how can the user or a program be given the information of what structures are worth considering? The answer is that the staircase algorithm may be given expert knowledge of the combinatorial relationships among the various Jordan structures known as *strata*. In this paper we discuss these relationships and complete the mathematical theory needed not only for the Jordan eigenvalue case, but also for the Jordan bundle (see section 2.3) problem, the Kronecker structure problem, and the Kronecker bundle problem.

Before we introduce the form of the relationships among the strata, it is helpful to think about the information produced from one iteration to the next in the staircase algorithm. Suppose that we already have determined a subblock corresponding to a single eigenvalue. For simplicity suppose that the eigenvalue is zero so that we know from the beginning that we are trying to find the Jordan structure of a nilpotent matrix. It is well known that the most *generic* such matrix has the single Jordan block $J_n(0)$ as its Jordan form. Such matrices form a dense set within the set of nilpotent matrices. Therefore all we know at first is that the matrix is in the closure of the matrices similar to a single Jordan block. As the staircase algorithm to deduce the Jordan form proceeds, we gain more and more information about the matrix. What in fact happens at each point in the algorithm is that we learn that the matrix is in the closure of the set of matrices similar to some other Jordan form. Indeed one may view the algorithm as identifying a nested sequence of closures. If during the course of the algorithm the user is unhappy with any choice, he or she might wish to have the power to backtrack through the algorithm and be offered other choices. Alternatively at the end of the algorithm the user might wish to know what he or she has missed in what has been described eloquently by Hough [30, p. 270] (in the polynomial case) as the "thicket" of nearby structures.

Following Arnold, Gusein-Zade, and Varchenko [3, p. 41] (also see historical and rather technical discussions by Goresky and MacPherson [27, pp. 33, 37]), we say that we have a *stratified manifold* if it is the union of nonintersecting manifolds whose closure is the finite union of itself with strata of smaller dimensions (thereby defining stratified manifolds recursively). For matrices, the strata are orbits of similar matrices, or perhaps the union of such orbits (known as *bundles*). For pencils, the strata consist of strictly equivalent pencils (or bundles). The problem of stratification is to find the closure relations among the various orbits or bundles. These relations define a partial ordering on orbits or bundles. One structure covers another if its closure includes the closure of the other and there is no other structure in between.

While we are the first to propose the use of stratifications in an algorithm, some of the mathematical theory, at least for nilpotent matrices, goes back to 1961. It is known to the lie algebra community as the closure ordering [11] and to the algebra community as degenerations of modules over the $\tilde{\mathbf{A}}_0$ quiver (see section 5.2) [9]. Combinatorically it is trivial; it is the well-known dominance ordering on partitions. This is the case of relevance in an algorithm when the eigenvalues are well clustered so that we may shift all the blocks to be nilpotent.

When eigenvalues are not well clustered, we have to consider the bundle case as defined by Arnold [2]. We have not seen the closure relation for this case in the literature so we believe that our theorems are new. We show that testing the closure relation for bundles leads to an NP-complete problem; therefore it may be expensive to use a stratification-enhanced algorithm in the bundle case when more than only a few eigenvalues need to be clustered.

For orbits of matrix pencils, the closure ordering was published in a linear algebra journal by Pokrzywa [37] in 1986. A general unifying algebraic theory of degenerations has been obtained for quivers by several authors including Abeasis and Del Fra [1] and Bongartz [9]. Bongartz studied the pencil case in 1990, apparently unaware of Pokrzywa's work. In algebraic language, for a matrix or pencil orbit or bundle to cover another, it is necessary to have an extension. This condition is not sufficient; another new result in this paper is the necessary and sufficient conditions for covers. The Kronecker bundle case also seems to be new.

We summarize the main theorems in the box below. The closure decision problem is the question of how to test whether the closure of a given orbit or bundle is contained within the closure of another. The closely related covering relationship tells us which structures are covered by a given structure. The * indicates that to the best of our knowledge the results were either previously unknown or not strong enough for purposes of numerical computations. (For example, we extend Pokrzywa's results by providing both necessary and sufficient conditions for one Kronecker orbit to cover another.)

	Closure Decision Problem	Covering Relationship
Jordan Orbits	Theorem 2.2	Corollary 2.3
Jordan Bundles	Theorem 2.6, 2.7 *	Theorem 2.6 *
Kronecker Orbits	Theorem 3.1	Theorem 3.2 *
Kronecker Bundles	Theorem 3.3 *	Theorem 3.3 *

We begin our paper reviewing the combinatorics of integer partitions in section 2.1. We then discuss the nilpotent orbit case already known in section 2.2, providing our own simple proof in terms of the staircase form. Section 2.3 addresses the bundle case showing that the decision procedure is an NP-complete problem. Section 3 covers the more complicated Kronecker case. In section 3.1 we state the stratification theorems for both orbits and bundles. Examples are worked out in section 3.2. Some special cases that arise in applications are further explored in sections 3.3 and 3.4. The proofs of the theorems may be found in section 3.5. Section 4 provides some of the necessary details for using the theorems inside of the staircase algorithms yielding our so-called "stratification-enhanced staircase algorithm." Finally, section 5 covers some mathematical aspects of the problem and also provides a short exposition on the algebraic notation so as to narrow the gap between the numerical and algebraic communities.

2. Stratification of the Jordan canonical form.

When the user of a numerical algorithm is confident in the clustering of the eigenvalues, then the only question that may arise is, What is the Jordan structure corresponding to an individual eigenvalue? In that case, there is no loss of generality assuming the eigenvalue is 0; hence we are interested in the stratification of orbits of nilpotent matrices, the topic of section 2.2. When we are less confident in the clustering, we must consider the stratification of bundles as discussed in section 2.3. We start with some elementary combinatorial notions.

2.1. Integer partitions.

A *partition* κ of an integer n is a sequence of integers $(k_1, k_2, k_3, \ldots)$ such that $k_1 + k_2 + \cdots = n$ and $k_1 \geq k_2 \geq \cdots \geq 0$. We use standard vector operations and if m is a scalar we denote $(k_1 + m, k_2 + m, \ldots)$ as $\kappa + m$. The partitions of an integer form a *lattice* ([40] is a good undergraduate reference) under the *dominance ordering*: the dominance ordering on partitions (or integer sequences) specifies that $\kappa \geq \lambda$ if and only if $k_1 + \cdots + k_i \geq l_1 + \cdots + l_i$, for $i = 1, 2, \ldots$, and we say that $\kappa > \lambda$ if and only if $\kappa \geq \lambda$ and $\kappa \neq \lambda$. To say that we have a lattice means that for every pair of partitions one can find an upper bound and a lower bound, i.e., a partition that dominates the pair, and a partition that is dominated by the pair. In a lattice we say that κ *covers* λ if and only if $\kappa > \lambda$, but there is no μ such that $\kappa > \mu > \lambda$. In Figure 2.1, the covering relationship for all integer partitions of $n = 8$ is illustrated in a *Hasse diagram.* Notice that we have placed the most dominant partition at the bottom of the diagram, i.e., the diagram shows the reversed dominance ordering.

```
    11111111
        |
     2111111
        |
     221111
   /       \
22211     311111
 |  \       /
2222  32111
 |    /  |
3221     |
 |    41111
3311     |
 |    \  |
332    4211
 |    /  |
422      |
 |     5111
431      |
 |  \    |
44     521
 |  /      \
53          611
    \      /
       62
        |
       71
        |
        8
```

FIG. 2.1. *Covering relationship for all integer partitions of $n = 8$.*

One can easily illustrate the covering relationship (Figure 2.2) by placing n coins in a table with k_1 in the first column, k_2 in the second column, etc., corresponding to a Ferrer diagram. A partition κ_1 covers κ_2 if κ_2 may be obtained from κ_1 by moving a coin rightward *one* column, or downward *one* row, so long as the partition remains monotonic [11]. Or equivalently, κ_1 covers κ_2 if κ_1 may be obtained from κ_2 by moving a coin leftward one column, or upward one row, and keeping the monotonicity of the partition. We call these moves a *minimum rightward* and a *minimum leftward coin move*, respectively.

The final elementary notion that we need is the *conjugate partition*, which is the partition obtained by "transposing" the coins and is here denoted κ'. Figure 2.3 shows how (3,2,2,1) and (4,3,1) are conjugate partitions. Since transposing reverses the direction of coin moves, it is clear that $\kappa > \lambda$ if and only if $\kappa' < \lambda'$.

2.2. Stratification of nilpotent orbits. Consider two set of matrices; the first consists of the matrices similar to the nilpotent matrix A_1 and the second is the set similar to the nilpotent matrix A_2. When is the closure of the second set similar to that of the first? The closure is a mathematically precise way to discuss the vague idea of a Jordan form being "near" another Jordan form.

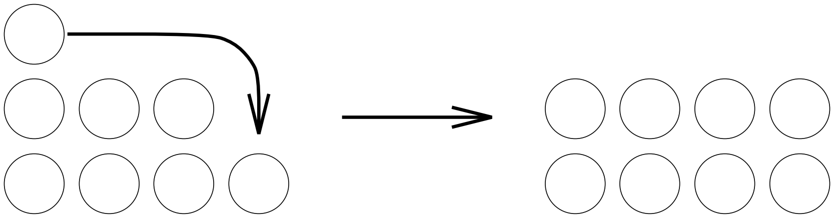


FIG. 2.2. *"Coin move" illustrates that* (3,2,2,1) *covers* (2,2,2,2).

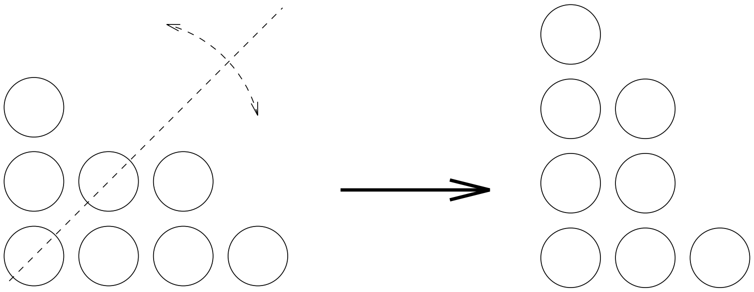


FIG. 2.3. *Transposing illustrates that* (3,2,2,1) *and* (4,3,1) *are conjugate partitions.*

Formally, in n^2-dimensional matrix space, consider the *orbits of matrices under similarity* transformations:

$$\mathcal{O}(A) \equiv \{SAS^{-1} : \det(S) \neq 0\}.$$

When is $\overline{\mathcal{O}}(A_1) \supseteq \overline{\mathcal{O}}(A_2)$? Trivially, if A_1 and A_2 are similar, then $\overline{\mathcal{O}}(A_1) = \overline{\mathcal{O}}(A_2)$. If $\overline{\mathcal{O}}(A_1) \supset \overline{\mathcal{O}}(A_2)$, then A_1 is "more generic" than A_2 or A_1 "degenerates" into A_2. In general, if an orbit $\mathcal{O}_1$ is more generic than an orbit $\mathcal{O}_2$, then $\dim \mathcal{O}_1 > \dim \mathcal{O}_2$. However, this is not a sufficient condition for the closure of $\mathcal{O}_2$ to be a proper subset of the closure of $\mathcal{O}_1$.

Associated with every nilpotent matrix A is the partition $\kappa(A) = (k_1, k_2, k_3, \ldots)$ that lists in decreasing order the sizes of the Jordan blocks associated with A. The k_i are known as the *Segre characteristics*. The partition $\mu(A)$ that is conjugate to $\kappa(A)$ contains what are known as the *Weyr characteristics*. (See, for example, [13] or older textbooks for discussion.) The *staircase form* [26, 33] is obtained by applying a unitary similarity transformation that puts the nilpotent matrix A in the form illustrated in Figure 2.4 for a partition with four parts. Here, the $A_{i,i+1}$ blocks are of full column rank, the *'s are arbitrary, the "lower staircase" (below the $A_{i,i+1}$ blocks) consists of only zero entries, and m_i (= the number of principal vectors of grade i) are the Weyr characteristics. The Weyr characteristics are the block sizes that appear in the staircase form. The nilpotent A in Figure 2.4 has $m_1 - m_2, m_2 - m_3, m_3 - m_4$, and m_4 Jordan blocks $J_i(0)$ of size $i = 1, 2, 3$, and 4, respectively.

Strictly upper triangular matrices are associated with directed acyclic graphs by taking the *sparsity graph*, meaning that node i points to node j if the (i, j) entry is not 0. Conversely, one can start with a directed acyclic graph G and find the Jordan structure of a generic matrix with sparsity graph G by a procedure suggested by Gansner [22]: a *path* in G is a sequence of vertices connected by directed edges in the usual orientation. A *k-path* is a subset of vertices that can be partitioned into

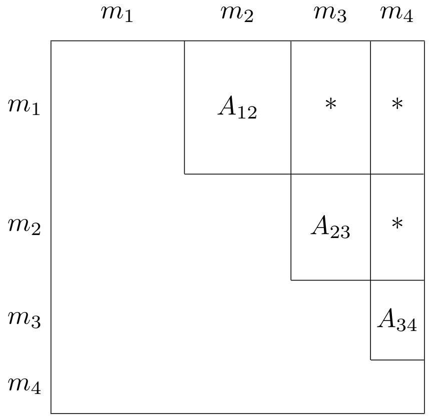


FIG. 2.4. *Example of a nilpotent A in staircase form.*

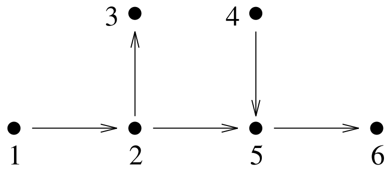


FIG. 2.5. *Digraph illustrating a $J_4 \oplus J_2$.*

k or fewer paths. Let s_1 denote the length of the longest path (1-path) in G, and inductively define s_j by letting $s_1 + \cdots + s_j$ denote the size of a longest (most vertices) j-path. These s_i are the Segre characteristics associated with the digraph. The dual notion to longest k-paths is the *shortest k-truncated path.* Consider a partition of the vertices of the digraph into paths labeled $1, \ldots, l$. Let w_i denote the length of the ith path or k, whichever is smaller. The length of such a k-truncated path is $w_1 + \cdots + w_l$; the smallest such sum is the length of the shortest k-truncated path. This gives a graph interpretation of the Weyr characteristics.

It would be a misconception that the size of the kth largest Jordan block can be found by looking at the longest path remaining after removing the longest $(k-1)$-path, since this may not be included in the longest k-path, as in the example in Figure 2.5. Here, the longest 1-path $(1, 2, 5, 6)$ of length four is not included in the longest 2-path $(1, 2, 3)$, $(4, 5, 6)$ of length six. (The Jordan normal form (JNF) of the generic matrix corresponding to the graph is $J_4 \oplus J_2$, not $J_4 \oplus J_1 \oplus J_1$.) By inspection we also see that this matrix cannot be put in staircase form by using only permutations. Even if the longest $(k-1)$-path is a subset of the longest k-path, it still may not necessarily be permuted into staircase form. The following is a graph characterization of the staircase form for nilpotent matrices.

THEOREM 2.1. *Denote the sources of a digraph as the* 1*-sources; deleting these sources we may denote the new sources as* 2*-sources, etc. A nilpotent matrix may be permuted to staircase form if and only if the k-sources form an antichain* (*i.e., no edges between them*) *and there is a matching between a subset of the* $(k-1)$*-sources and the k-sources.*

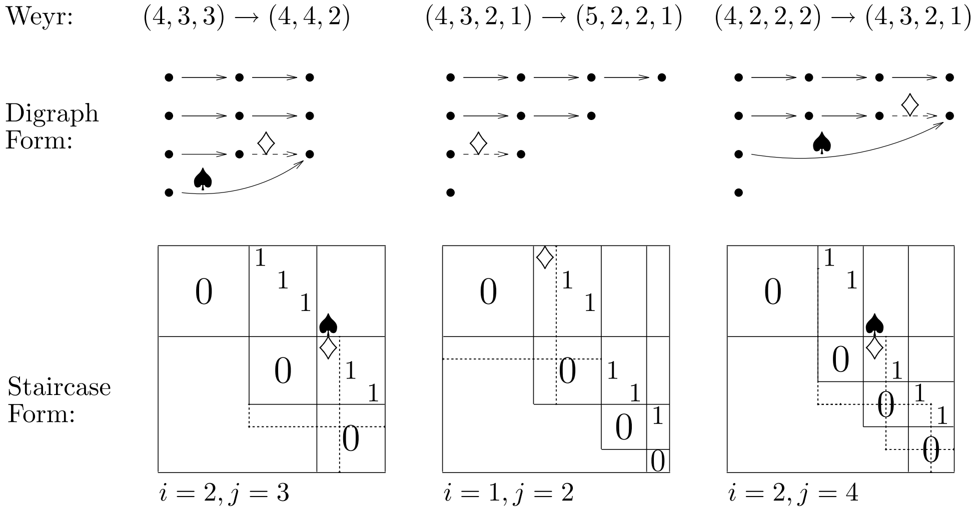


FIG. 2.6. *The* $\diamondsuit$ *goes from* 1 *to* 0. *The* $\spadesuit = 1$ *is introduced to preserve rank conditions. The solid lines indicate the original Jordan structure in staircase form. The dashed lines indicate the diagonal blocks of the final Jordan structure obtained when* $\diamondsuit = 0$.

Proof. The nodes in the k-sources correspond to the kth diagonal block of the staircase form. The antichain corresponds to $A_{k,k} = 0$, and the matching condition is the full-rank condition on $A_{k,k+1}$. ☐

Our own version of the proof of the nilpotent stratification is quite short. Much of the proof may be understood by inspection of the staircase form or the digraph.

THEOREM 2.2. *$\overline{\mathcal{O}}(A_1) \supseteq \overline{\mathcal{O}}(A_2)$ if and only if $\mu(A_2) \geq \mu(A_1)$, or equivalently $\kappa(A_1) \geq \kappa(A_2)$, where μ and κ denote the Weyr and Segre characteristics, respectively.*

Proof. We first remark that if $A_2 \in \overline{\mathcal{O}}(A_1)$, then $\mathcal{O}(A_2) \subseteq \overline{\mathcal{O}}(A_1)$ by taking similarity transformations. It then follows upon taking the closure of both sides that $\overline{\mathcal{O}}(A_2) \subseteq \overline{\mathcal{O}}(A_1)$. Therefore, to prove the "if" assertion, it suffices to assume that μ_2 covers μ_1 and to exhibit an A_1 and A_2 such that $\mu(A_i) = \mu_i$ for $i = 1, 2$ and $A_2 \in \overline{\mathcal{O}}(A_1)$.

For any μ_2 that covers μ_1, we have a "coin move" that decrements a column of size m_j and increments that of m_i, where $i < j$. The A_1 that we will pick with $\mu(A_1) = \mu_1$ has square identity matrices placed at the top of the superdiagonal blocks, except for $A_{j,j+1}$, where the identity matrix is placed at the bottom (if $j < j_{\max}$, the size of the largest Jordan block). We also place a 1 (denoted $\spadesuit$ in Figure 2.6) in the first column of the super-super diagonal block $A_{i-1,i+1}$ in the last row (if $i > 1$). By zeroing the 1 in the first column of $A_{i,i+1}$ continuously, we effect a coin move that reduces m_j by one and increases m_i by one.

In Figure 2.6 we illustrate a few cases. The diamond ($\diamondsuit$) moves continuously from 1 to 0. The spade ($\spadesuit = 1$) is introduced if $i > 1$. Note the cascading effect when $j \neq i + 1$ in the equal blocks. The graph picture of the proof in terms of paths is also displayed in Figure 2.6. The deletion of the edge with the diamond corresponds to a coin move.

To prove the "only if" assertion, we assume that A_2 is a limit point of a continuous path Γ in $\mathcal{O}(A_1)$. We may continuously decompose every point on Γ into the staircase form corresponding to the partition $\mu(A_1)$. By the boundedness of the set of orthogonal matrices (compactness of the orthogonal group), although A_2, which is in the closure of $\mathcal{O}(A_1)$, may not be in $\mathcal{O}(A_1)$, it may also be put into the very same staircase form, although we may lose the full rank conditions on the $A_{i,i+1}$.

A rank one drop in one such $A_{i,i+1}$ corresponds to a leftward coin move. For example, if the rank of A_{23} in Figure 2.4 drops by one, then after an orthogonal similarity transformation, m_3 is decremented by one and m_2 (or possibly m_1 if $m_1 = m_2$) increments by one. (This assumes that the matrix consisting of A_{13} on top of A_{23} does not itself lose column rank, otherwise we have a more "long-distance" coin move obtainable by cascading short coin moves.) Therefore $\mu(A_2)$ dominates $\mu(A_1)$. $\square$

From Theorem 2.2 and the definition of covering partitions we get the following obvious characterization for covering orbits of nilpotent matrices A_1 and A_2.

COROLLARY 2.3. *$\mathcal{O}(A_1)$ covers $\mathcal{O}(A_2)$ if and only if $\mu(A_2)$ can be obtained from $\mu(A_1)$ by a minimum leftward coin move.*

By reading the Hasse diagram in Figure 2.1 from top to bottom we get the stratification in terms of the Weyr characteristics. Reading the diagram from bottom to top, we get the closure hierarchy in terms of the Segre characteristics.

2.3. Stratification of Jordan bundles.

Let $J_n(\alpha)$ denote a single $n \times n$ Jordan block with eigenvalue α. Our first example of a *bundle* as defined by Arnold [2, Sect. 5.3] is $\bigcup_\alpha \mathcal{O}(J_n(\alpha))$, the set of all matrices whose Jordan form consists of a single block. Notice that the bundle is the union of orbits. Here is the general definition. If two matrices have the same Jordan structure except that the distinct eigenvalues are different, we say they are in the same bundle. More precisely, let $w(\lambda_1), w(\lambda_2), \ldots, w(\lambda_p)$ be the Weyr characteristics of a matrix A with distinct eigenvalues $\lambda_1, \ldots, \lambda_p$. (Remember that $w(\lambda_i)$ is a partition of n_i, the algebraic multiplicity of the eigenvalue λ_i.) Another matrix B is said to be in the bundle $\mathcal{B}(A)$ if the distinct eigenvalues $\mu_1, \ldots, \mu_p$ of B may be ordered in such a way that the sequence of partitions $w(\mu_1), w(\mu_2), \ldots, w(\mu_p)$ is identical to that of A.

Let A_1 and A_2 be two nilpotent matrices. We define $A_1 \boxplus A_2$ to be the matrix in Jordan form (with the Jordan blocks ordered in decreasing order) whose Segre characteristics are the sums of those of A_1 and A_2, or equivalently whose Weyr characteristics are the union of those of A_1 and A_2. We point out that $A_1 \oplus A_2$ goes the other way: the Segre characteristics are the union, and the Weyr characteristics are the sum. For example, if $A_1 = J_3(0) \oplus J_1(0)$ and $A_2 = J_3(0) \oplus J_2(0) \oplus J_1(0)$, then $A_1 \oplus A_2 = 2J_3(0) \oplus J_2(0) \oplus 2J_1(0)$ and $A_1 \boxplus A_2 = J_6(0) \oplus J_3(0) \oplus J_1(0)$. We define an *extension* of A_1 and A_2 to be any matrix of the form

$$\begin{pmatrix} A_1 & X \\ 0 & A_2 \end{pmatrix},$$

where X could be any matrix of conforming size. For the example above, we get an extension with Jordan structure $A_1 \boxplus A_2$ by choosing x_{31} and x_{44} nonzero and all other elements in X (of size 4×6) as zero.

We have already pointed out that when $X = 0$, the Segre characteristics of the extension is the union of the Segre characteristics of A_1 and A_2. Therefore, an easy consequence of Theorem 2.2 is that $\mathcal{O}(A_1 \oplus A_2) \subset \overline{\mathcal{O}}(A_1 \boxplus A_2)$.

LEMMA 2.4. *The most generic extensions of A_1 and A_2 are in the orbit of $A_1 \boxplus A_2$.*

Proof. The easiest proof of this statement is obtained by assuming that A_1 and

A_2 are in Jordan form and by examining the longest k-paths in an extension. At most we can connect the longest path from the graph of A_1 to that of A_2, then the next longest of each, etc. Another proof may also be found in [25, Prop. 4.2.2]. ∎

If A is an extension of A_1 and A_2, then we have the obvious statement that

$$\mathcal{O}(A_1 \oplus A_2) \subseteq \overline{\mathcal{O}}(A) \subseteq \overline{\mathcal{O}}(A_1 \boxplus A_2). \tag{2.1}$$

The set of matrices A satisfying the relation (2.1) form a sublattice of the dominance ordering. Unfortunately, in general, this sublattice is not the set of extensions of A_1 and A_2. (An example is $A_1 = J_6(0)$ and $A_2 = J_4(0) \oplus J_2(0)$. In the lattice, $J_6(0) \oplus J_5(0) \oplus J_1(0)$ is between $A_1 \boxplus A_2$ and $A_1 \oplus A_2$, but it is not an extension of A_1 and A_2.) The characterization of the extensions (a further sublattice of this sublattice) is an open problem according to [25, p. 133], but it is not needed for our purposes.[1]

In the next lemma, we consider limit points of continuous paths $A(t)$ such that when $0 \le t < 1$, the path is contained in a bundle consisting of two distinct eigenvalues.

LEMMA 2.5. *Suppose* $A(t)$ *is similar to* $A_1(t) \oplus A_2(t)$ *for* $0 \le t \le 1$*, where* $A_1(t) - \beta(t)I$ *and* $A_2(t) - \gamma(t)I$ *are nilpotent, and for* $0 \le t < 1$*,* $\beta(t) \ne \gamma(t)$*, but when* $t = 1$*,* $\beta(1) = \gamma(1) = 0$*. In other words,* $A_1(t)$ *has the unique eigenvalue* $\beta(t)$*,* $A_2(t)$ *has the unique eigenvalue* $\gamma(t)$*, and these eigenvalues coalesce at* 0 *when* $t = 1$*. Then*

$$\overline{\mathcal{O}}(A(1)) \subseteq \overline{\mathcal{O}}(A_1(1) \boxplus A_2(1)).$$

Proof. We may find a continuous orthogonal similarity transformation $Q(t)$ such that

$$Q^T(t)A(t)Q(t) = \begin{pmatrix} A_1'(t) & X(t) \\ 0 & A_2'(t) \end{pmatrix},$$

and $A_1'(t)$ is similar to $A_1(t)$ and $A_2'(t)$ is similar to $A_2(t)$ for $0 \le t < 1$. Therefore, by Lemma 2.4,

$$\begin{pmatrix} A_1'(t) & X(t) \\ 0 & A_2'(t) \end{pmatrix} - \begin{pmatrix} \beta(t)I & \\ & \gamma(t)I \end{pmatrix} \in \overline{\mathcal{O}}(A_1(1) \boxplus A_2(1))$$

for $0 \le t < 1$. Letting $t \to 1$ shows that $A(1) \in \overline{\mathcal{O}}(A_1(1) \boxplus A_2(1))$, from which the result follows. ∎

The Jordan bundle stratification theorem follows below. Our results for the closure decision problem are also derived in [36, 17]. We believe the covering relationship is new.

THEOREM 2.6. *Suppose that we have two bundles* $\mathcal{B}(A_1)$ *and* $\mathcal{B}(A_2)$ *such that the former has at least as many distinct eigenvalues as the latter. Then* $\overline{\mathcal{B}}(A_1) \supseteq \overline{\mathcal{B}}(A_2)$ *if and only if it is possible to coalesce eigenvalues and apply the dominance ordering coin moves to the bundle defined by* A_1 *to reach that of* A_2*. Furthermore, a cover is obtained either by a minimal coin move* (*on the structure for one eigenvalue*) *or a generic extension* (*of the structures for two distinct eigenvalues assumed to coalesce*)*.*

Proof. All that remains is to prove the minimality property of these covering relations. There are two natural quotient lattices of the bundle lattice. The first

---

[1] We prefer the use of the word "extension" rather than "completion" as used by [25] for consistency with the algebraic notion of extension of two modules.

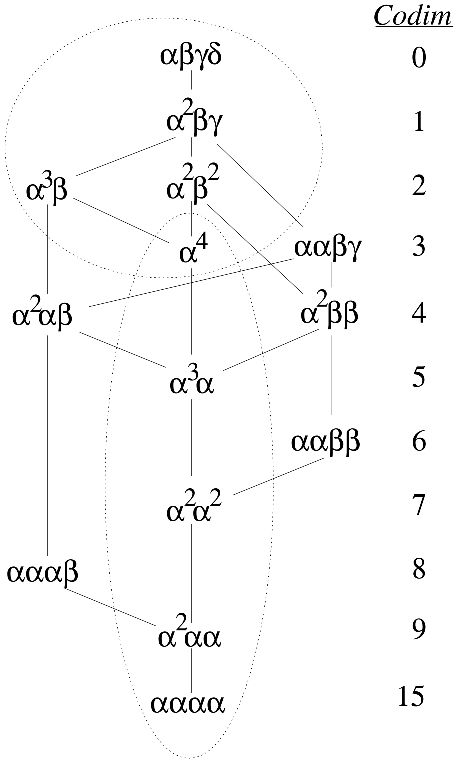


FIG. 2.7. *The bundle stratification for* 4×4 *matrices.*

counts the number of eigenvalues. The second is the partition of n obtained by taking the union of the partitions corresponding to all of the eigenvalues. Moving a coin in one partition moves down the first lattice but not the second. Coalescing two eigenvalues (forming the union of two partitions) moves down the second lattice but not the first. Therefore each operation cannot be obtained from the result of the other operation, so each is minimal. □

Figure 2.7 plots the bundle stratification for 4×4 matrices. Here, we use Arnold's compact notation for Jordan blocks: $\alpha^k \equiv J_k(\alpha)$. Circled in the top of the figure are those structures corresponding to coalescing eigenvalues. It is possible to gain an appreciation of the complicated manner of how these structures fit inside each other from the swallowtail diagram in Figure 2.8 [2], which shows the projection of these structures into three-dimensional space. The point α^4 is the swallowtail point, the curve $\alpha^3\beta$ are the two cusp edges coming out from the swallowtail, $\alpha^2\beta^2$ is the transversal intersection of the wings, and $\alpha^2\beta\gamma$ is the surface of the swallowtail. Everything outside the swallowtail is represented by $\alpha\beta\gamma\delta$. Any reliable numerical attempt to find the nearest structure of a certain particular form must somehow implicitly or explicitly deal with this kind of geometry. The circled structures in the lower part of the figure are those that correspond to the stratification of nilpotent orbits.

Figure 2.7 captures all distinct singularities of codimension 1 (α^2) and 2 (α^3, $\alpha^2\beta^2$) and two of the four distinct bundles of codimension 3 (α^4 and $\alpha\alpha$). The missing ones are $\alpha^3\beta^2$ and $\alpha^2\beta^2\gamma^2$.

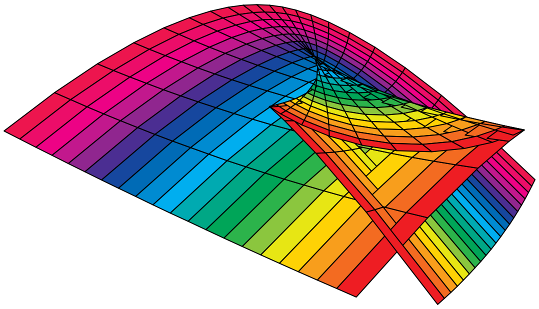


FIG. 2.8. *The swallowtail.*

Unfortunately, although the decision procedure for testing the closure relation for nilpotent matrices is trivial (all that is required is to test if one partition dominates the other), the corresponding procedure for bundles is an NP-complete problem. We speculate that this may capture some of the essence of the true difficulty associated with the clustering problem for perturbed eigenvalues. Another result that is slightly related was obtained by Gu [28], who showed that finding a well-conditioned similarity transformation to block-diagonalize a nonsymmetric matrix is an NP-hard problem.

THEOREM 2.7. *Deciding whether a bundle is in the closure of another bundle is an NP-complete problem.*

Proof. Suppose that we have a matrix of dimension $n = km$ that has $3m$ distinct eigenvalues with multiplicities $k_1, \ldots, k_{3m}$ with the property that $k/4 < k_i < k/2$ for each i. Consider the existence of a clustering of all of these eigenvalues into m triples so that the sum of the multiplicities of the three eigenvalues in each cluster is exactly k. This problem is the three-partition problem and is well known to be NP-complete [24]. ☐

The implication of Theorem 2.7 is that it is unlikely to find an efficient algorithm that solves all instances of the decision problem. However, it is still possible that there exist algorithms that can solve most practical cases efficiently.

3. Stratification of the Kronecker canonical form.

The notions of canonical form, orbits, bundles, and partitions extend to the matrix pencil case in a straightforward manner as follows. Any matrix pair (A, B), where A and B are $m \times n$ with real or complex entries, defines an *orbit (manifold) of strictly equivalent matrix pencils* in the $2mn$-dimensional space $\mathcal{P}$ of $m \times n$ pencils:

$$\mathcal{O}(A - \lambda B) = \{U^{-1}(A - \lambda B)V : \det(U)\det(V) \neq 0\}. \tag{3.1}$$

Let $P_1 = A_1 - \lambda B_1$, $P_2 = A_2 - \lambda B_2$ be two pencils (of possibly different sizes). A pair (U, V) satisfying $UP_1 = P_2 V$ defines a homomorphism from P_1 to P_2 (see section 5.1). Let the dimension of such (U, V) be

$$\langle P_1, P_2 \rangle = \dim\{(U, V) : UP_1 = P_2 V\}. \tag{3.2}$$

We also define $\mathrm{cod}(A - \lambda B)$ as the *codimension* of $\mathcal{O}(A - \lambda B)$, which is the dimension of the space complementary to the orbit, e.g., the dimension of the space normal to $\mathcal{O}(A - \lambda B)$ at the point $A - \lambda B$ [43, 13, 19]. It is known that $\mathrm{cod}(P) = \langle P, P \rangle - (m-n)^2$.

The *Kronecker canonical form* (KCF) (e.g., see [23]) for a pencil is the direct sum of the right singular, left singular, and regular structures, consisting of L_k blocks of size $k \times (k+1)$ for the *right singular structure* and L_k^T blocks for the *left singular structure*. The *regular structure* consists of Jordan blocks $J_k(\mu)$ corresponding to eigenvalue μ and N_k for the eigenvalue ∞. For short, we omit the word "singular" when clear from context. If $A - \lambda B$ is $m \times n$, where $m \neq n$, then for almost all A and B it will have the same KCF, depending only on m and n (the *generic case*; see section 3.3).

Define the *normal rank* of $A - \lambda B$ as

$$\mathrm{nrk}(A - \lambda B) = n - r_0 = m - l_0,$$

where r_0 and l_0 are the total number of right and left blocks, respectively, in the KCF of $A - \lambda B$.

As in the matrix case we consider two main problems in order to understand the stratification of matrix pencils. First, given two $m \times n$ matrix pencils $P_1 = A_1 - \lambda B_1$ and $P_2 = A_2 - \lambda B_2$ we want to have a procedure for deciding whether $\overline{\mathcal{O}}(P_1) \supseteq \overline{\mathcal{O}}(P_2)$. Second, we want to find a procedure for generating covering pencils, i.e., the closest neighboring orbits above or below in the closure hierarchy. These two problems are also investigated for bundles of pencils.

Because of the three basic structures (right, left, regular) in the Kronecker form, the complete characterization of all possible $m \times n$ Kronecker forms, their orbits, and bundles is somewhat more intricate than for the matrix case. We not only will need to define partitions for each eigenvalue, but in addition we will need partitions to define the right and left singular structures.

Let $\mathcal{R}(P)$ and $\mathcal{L}(P)$ denote the partitions for the right and left structures, respectively, of $P = A - \lambda B$ and let $\mathcal{J}_\mu(P)$ denote the partition for the Jordan structure corresponding to the eigenvalue μ (finite or infinite). When it is clear from context, we use the abbreviated notation $\mathcal{R}$, $\mathcal{L}$, and $\mathcal{J}_\mu$. The j_i's in a $\mathcal{J}$ partition are the Weyr characteristics for the eigenvalue μ, i.e., j_i is the number of $J_k(\mu)$ blocks of size $k \geq i$. Similarly, r_i of $\mathcal{R}$ (or l_i of $\mathcal{L}$) is the number of L_k (or L_k^T) blocks of size $k \geq i$.

We have shown in section 2 that the closure hierarchy of the set of $n \times n$ nilpotent matrices is completely determined by the dominance ordering of the integer n. Since the Kronecker structure for matrix pencils includes both Jordan blocks and singular blocks, a corresponding characterization involves integer sequences corresponding to each kind of block.

3.1. Stratification of Kronecker orbits and bundles. The decision procedure for the closure of orbits was derived by Pokrzywa [37] in 1986, was later reformulated by De Hoyos [12] in 1990, and was formulated differently by Bongartz [9] in 1990. In the following we give our formulation of De Hoyos's closure characterization. In section 5.2 we explain the algebraic and geometric connections between these approaches.

THEOREM 3.1 (see [37, 12, 9]). *$\overline{\mathcal{O}}(P_1) \supseteq \overline{\mathcal{O}}(P_2)$ if and only if the following relations hold:*

- $\mathcal{R}(P_1) + \mathrm{nrk}(P_1) \geq \mathcal{R}(P_2) + \mathrm{nrk}(P_2)$,
- $\mathcal{L}(P_1) + \mathrm{nrk}(P_1) \geq \mathcal{L}(P_2) + \mathrm{nrk}(P_2)$,
- $\mathcal{J}_{\mu_i}(P_1) + r_0(P_1) \leq \mathcal{J}_{\mu_i}(P_2) + r_0(P_2)$

for all $\mu_i \in \overline{\mathbf{C}}$, $i = 1, 2, \ldots$, where $\overline{\mathbf{C}} = \mathbf{C} \bigcup \{\infty\}$.

We remark that we could have used the more symmetric looking expression $\mathcal{J}_{\mu_i}(P_1) - \mathrm{nrk}(P_1) \leq \mathcal{J}_{\mu_i}(P_2) - \mathrm{nrk}(P_2)$ as item three in the theorem at the cost of having "negative coins." If $\mathrm{nrk}(P_1) = \mathrm{nrk}(P_2)$, we say that P_1 and P_2 are on the same level playing field, and the relations in Theorem 3.1 reduce to $\mathcal{R}(P_1) \geq \mathcal{R}(P_2)$, $\mathcal{L}(P_1) \geq \mathcal{L}(P_2)$, and $\mathcal{J}_{\mu_i}(P_1) \leq \mathcal{J}_{\mu_i}(P_2)$.

In a contemporary paper [7], Boley applies Theorem 3.1 and shows similar majorizing results of integer sequences associated with the KCF when a single row or column is appended to a matrix pencil. The application considered is adding a single input or output to a linear time-invariant dynamical system.

Now that we can test if one structure is more generic than another in the closure lattice, the next question we consider is the generation of all structures covered by a given pencil. Necessary conditions for two structures to be (closest) neighbors in the lattice are given in [37, 12, 9] and are used in an algorithm for computing the complete Kronecker structure hierarchy in [20]. We believe we are the first to give both necessary and sufficient conditions for neighbors in the lattice (which in addition gives an optimal algorithm for computing the complete hierarchy). We present these results in the form of coin moves associated with $\mathcal{R}$, $\mathcal{L}$, and $\mathcal{J}_{\mu_i}$ for different eigenvalues μ_i.

THEOREM 3.2. *$\mathcal{O}(P_1)$ covers $\mathcal{O}(P_2)$ if and only if P_2 can be obtained by applying one of the rules* (1)–(4) *to the integer partitions of P_1:*

(1) *Minimum rightward coin move in $\mathcal{R}$ (or $\mathcal{L}$).*

(2) *If the rightmost column in $\mathcal{R}$ (or $\mathcal{L}$) is one single coin, append that coin as a new rightmost column of some $\mathcal{J}_{\mu_i}$ (which may be empty initially).*

(3) *Minimum leftward coin move in any $\mathcal{J}_{\mu_i}$.*

(4) *Let k denote the total number of coins in all of the longest ($=$ lowest) rows from all of the $\mathcal{J}_{\mu_i}$. Remove these k coins, add one more coin to the set, and distribute $k+1$ coins to r_p, $p = 0, \ldots, t$ and l_q, $q = 0, \ldots, k-t-1$ such that at least all nonzero columns of $\mathcal{R}$ and $\mathcal{L}$ are given coins.*

Rules (1) *and* (2) *may not make coin moves that affect r_0 (or l_0).*

Notice that the restriction for rules (1) and (2) implies that the number of left and right blocks remains fixed, while rule (4) adds one new block of each kind. We also remark that rule (4) cannot be applied if the total number of nonzero columns in $\mathcal{R}$ and $\mathcal{L}$ is more than $k+1$. Rule (3) corresponds to the nilpotent case.

As in the matrix case we also consider stratification of bundles. Two pencils are in the same bundle if they have the same left and right singular structures and the same Jordan structure except that the distinct eigenvalues may be different. Of course, if $\mathcal{O}(P_1)$ covers $\mathcal{O}(P_2)$, then $\overline{\mathcal{B}}(P_1) \supset \overline{\mathcal{B}}(P_2)$, but the two bundles may not necessarily be covering, since there is a possibility of other structure changes from coalescing eigenvalues.

THEOREM 3.3. *$\mathcal{B}(P_1)$ covers $\mathcal{B}(P_2)$ if and only if P_2 can be obtained by applying one of the rules* (1)–(5) *to the integer partitions of P_1:*

(1) *Same as rule* 1 *in Theorem* 3.2.

(2) *Same as rule* 2 *in Theorem* 3.2, *except it is allowed only to start a new set corresponding to a new eigenvalue (i.e., no appending to nonempty sets).*

(3) *Same as rule* 3 *in Theorem* 3.2.

(4) *Same as rule* 4 *in Theorem* 3.2, *but apply only if there is just one eigenvalue in the KCF or if all eigenvalues have at least two Jordan blocks.*

(5) *Let any pair of eigenvalues coalesce, i.e., take the union of their sets of coins.*

The problem of deciding if the closure of the bundle of one pencil contains the bundle of another is NP-complete, just as for the matrix case (see Theorem 2.7). We postpone the proofs of Theorems 3.2 and 3.3 to section 3.5 and continue by illustrating the theorems with some examples. In section 4 an algorithmic implication of Theorem 3.2 is presented. We express different structure transitions in rules (1)–(4) in terms of the structure indices computed by a staircase algorithm.

$$P_1 = L_0 \oplus L_1 \oplus L_2 \oplus J_1(\mu_1) \oplus J_4(\mu_1) \oplus J_3(\mu_2) \oplus L_1^T \oplus L_3^T$$

	Structure blocks	Partition	Coins
Right	$L_0 \oplus L_1 \oplus L_2$	$\mathcal{R} = (3, 2, 1)$	○ ○ ○ ○ ○ ○
Left	$L_1^T \oplus L_3^T$	$\mathcal{L} = (2, 2, 1, 1)$	○ ○ ○ ○ ○ ○
Regular eig=μ_1	$J_1(\mu_1) \oplus J_4(\mu_1)$	$\mathcal{J}_{\mu_1} = (2, 1, 1, 1)$	○ ○ ○ ○ ○
eig=μ_2	$J_3(\mu_2)$	$\mathcal{J}_{\mu_2} = (1, 1, 1)$	○ ○ ○

FIG. 3.1. *Example Kronecker structure P_1 with corresponding partitions.*

3.2. Examples. To focus the reader's attention on how the covering theorem may be used in a numerical algorithm, we will examine two examples in detail. In the first example we take a particular pencil P_1 of size 17×18 and illustrate the application of some of the rules from Theorem 3.2. In the second example, we focus on a smaller case—the 2×3 pencils—and show the entire lattices for orbits and bundles. We also use this example to illustrate Theorem 3.1.

Our example pencil P_1 has KCF $L_0 \oplus L_1 \oplus L_2 \oplus J_1(\mu_1) \oplus J_4(\mu_1) \oplus J_3(\mu_2) \oplus L_1^T \oplus L_3^T$. We illustrate how the four rules in Theorem 3.2 can be used to find a pencil P_2 that is covered by P_1. The starting configuration for P_1 may be found in Figure 3.1 and the application of some of the rules is shown in Figure 3.2. We display KCF structures as both integer partitions and columns of "coins" ($\circ$). In Figure 3.2 we illustrate how each of the rules can be applied to P_1. (We append the notation a and b to rules (1) and (2) to denote application to the right or left structure, respectively.) The symbol $\bullet$ is used to denote a coin that will be moved in P_1's coin arrays or a coin that has been moved in P_2's coin arrays. Notice that some of the rules can also be applied to other combinations of blocks of P_1, i.e., the figure does not show all possible transitions that give a pencil P_2 that is covered by P_1. Each row of the figure shows how one of the rules may be applied to some of the blocks in the KCF. In the last column of Figure 3.2 (labeled "Block transitions") we record only the blocks that are involved in the application of the rule.

Recently, Elmroth and Kågström did a comprehensive study of the set of 2×3 pencils, including the stratification problem [21]. There are 9 possible bundles in this case. (From an algorithmic point of view, we may not want to bundle in the zero and infinite eigenvalue, in which case there are 20 bundles.) Fix the eigenvalues in the bundle to be γ and δ, with $\gamma \neq \delta$. The closure lattice corresponding to the orbits is shown in Figure 3.3. Following [21] we display the lattice with orbits (nodes) of the same codimension on the same horizontal level. The generic case (L_2) is at the highest level and the most degenerate pencil ($3L_0 \oplus 2L_0^T$ which is the 2×3 zero pencil) is at the lowest level. A pencil P_2 is in $\overline{\mathcal{O}}(P_1)$ if and only if there is a path from P_1 to P_2. The labels of the arcs correspond to which covering rules in Theorem 3.2 we have applied.

In Figure 3.4 we show the closure lattice corresponding to the bundles of 2×3 pencils. Here, a node represents the bundle consisting of all pencils with the displayed Kronecker structure, where the value of the eigenvalues γ and δ may vary, but $\gamma \neq \delta$. The arc labels show which of the rules in Theorem 3.3 we have applied. Since the

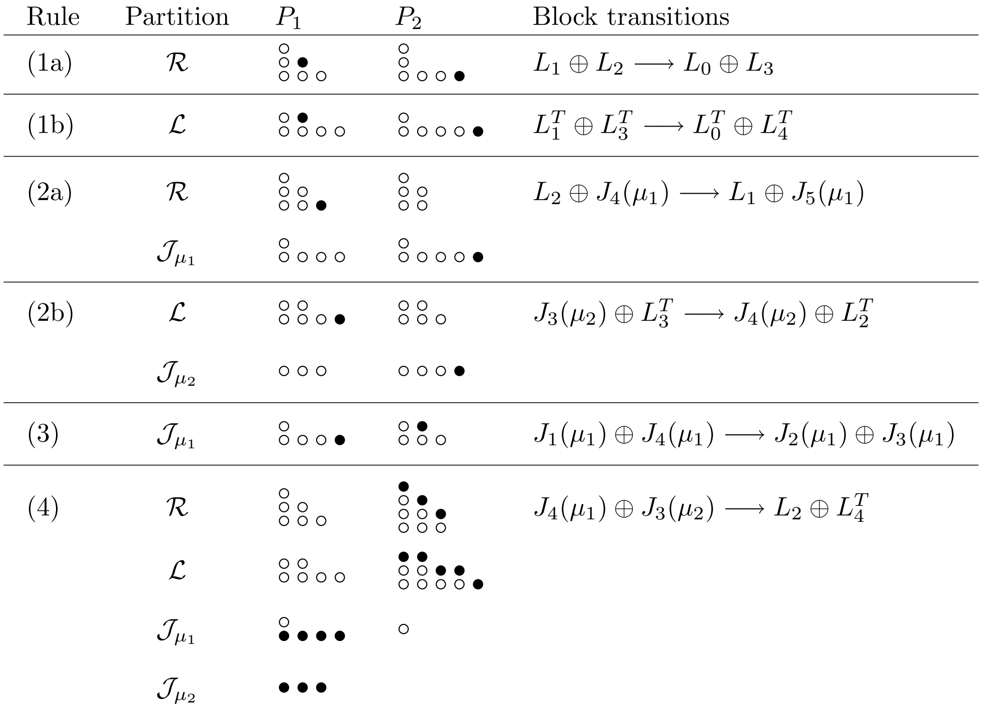

Rule
Partition
P_1
P_2
Block transitions
(1a)
$\mathcal{R}$
$L_1 \oplus L_2 \longrightarrow L_0 \oplus L_3$
(1b)
$\mathcal{L}$
$L_1^T \oplus L_3^T \longrightarrow L_0^T \oplus L_4^T$
(2a)
$\mathcal{R}$
$L_2 \oplus J_4(\mu_1) \longrightarrow L_1 \oplus J_5(\mu_1)$
$\mathcal{J}_{\mu_1}$
(2b)
$\mathcal{L}$
$J_3(\mu_2) \oplus L_3^T \longrightarrow J_4(\mu_2) \oplus L_2^T$
$\mathcal{J}_{\mu_2}$
(3)
$\mathcal{J}_{\mu_1}$
$J_1(\mu_1) \oplus J_4(\mu_1) \longrightarrow J_2(\mu_1) \oplus J_3(\mu_1)$
(4)
$\mathcal{R}$
$J_4(\mu_1) \oplus J_3(\mu_2) \longrightarrow L_2 \oplus L_4^T$
$\mathcal{L}$
$\mathcal{J}_{\mu_1}$
$\mathcal{J}_{\mu_2}$

FIG. 3.2. *Illustration of the covering rules in Theorem* 3.2 *starting with* P_1 *as in Figure* 3.1.

eigenvalues may vary in the bundles, the codimension for each Kronecker structure with regular part is one less for each eigenvalue compared to the orbit case (see Figure 3.3). For example, the codimension of $L_1 \oplus J_1(\gamma)$ is 2 in Figure 3.3 and 1 in Figure 3.4, since γ is an extra degree of freedom in the bundle case.

When comparing the closure hierarchies for bundles and orbits, we see that the bundle structure $L_0 \oplus J_2(\gamma)$ is found as the most generic one when γ and δ coalesce (rule (5)) in $L_0 \oplus J_1(\gamma) \oplus J_1(\delta)$, while these two structures are on different branches in the hierarchy for orbits (where the eigenvalues are assumed to be specified and therefore never may coalesce). This illustrates the restriction of rule (2) in Theorem 3.3.

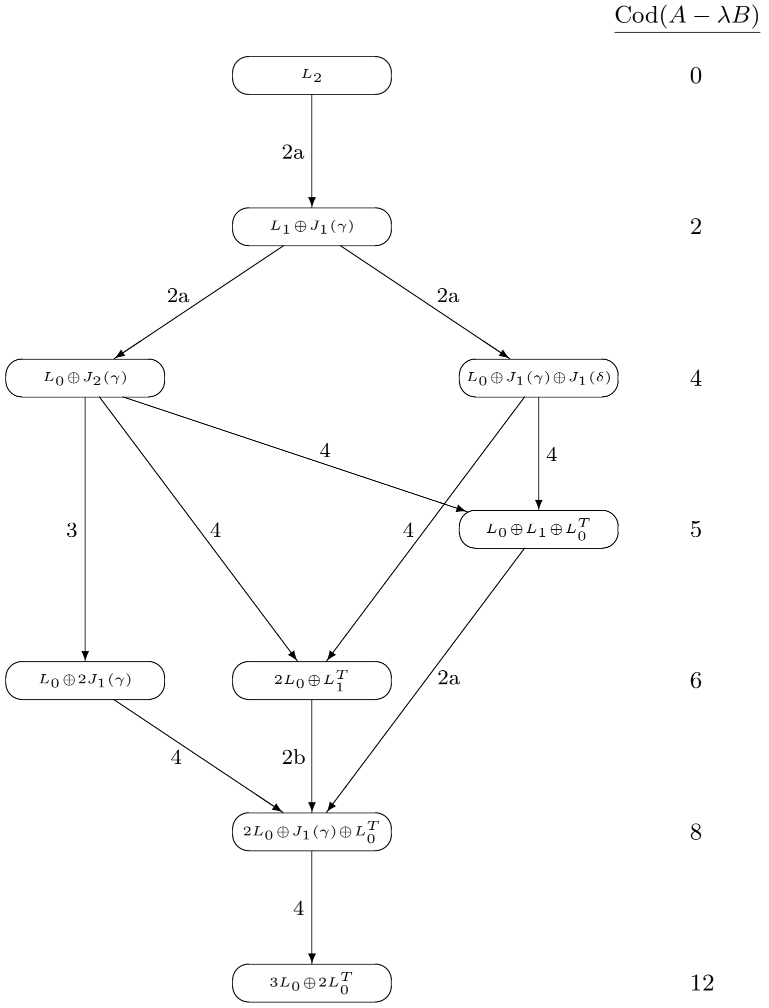
Finally, we illustrate Theorem 3.1 by investigating the closure relations for the orbits of 2×3 pencils $P_1 = L_1 \oplus J_1(\gamma)$, $P_2 = L_0 \oplus J_2(\gamma)$, $P_3 = L_0 \oplus J_1(\gamma) \oplus J_1(\delta)$, and $P_4 = 2L_0 \oplus L_1^T$. From Table 3.1 we see that P_2, P_3, P_4 are in $\overline{\mathcal{O}}(P_1)$, P_4 is in P_1, P_2 and P_3, but neither P_2 or P_3 are in the closure of the other. To realize that P_1 and P_4 are closest neighbors to both P_2 and P_3 we have to apply Theorem 3.2 (see Figure 3.3).

3.3. Generic and full normal rank pencils.

The *generic Kronecker structure* for $A - \lambda B$ of size $m \times n$ with $d = n - m > 0$ is

$$\operatorname{diag}(L_\alpha, \ldots, L_\alpha, L_{\alpha+1}, \ldots, L_{\alpha+1}), \tag{3.3}$$

where $\alpha = \lfloor m/d \rfloor$, the total number of blocks is d, and the number of $L_{\alpha+1}$ blocks is m mod d (which is 0 when d divides m) [41, 13]. The same statement holds for $d = m - n > 0$ if we replace $L_\alpha, L_{\alpha+1}$ in (3.3) by $L_\alpha^T, L_{\alpha+1}^T$. Indeed, a generic



$\mathrm{Cod}(A - \lambda B)$
L_2
0
2a
$L_1 \oplus J_1(\gamma)$
2
2a
2a
$L_0 \oplus J_2(\gamma)$
$L_0 \oplus J_1(\gamma) \oplus J_1(\delta)$
4
4
4
3
4
4
$L_0 \oplus L_1 \oplus L_0^T$
5
$L_0 \oplus 2J_1(\gamma)$
$2L_0 \oplus L_1^T$
2a
6
4
2b
$2L_0 \oplus J_1(\gamma) \oplus L_0^T$
8
4
$3L_0 \oplus 2L_0^T$
12

FIG. 3.3. *Closure hierarchy of orbits for the set of* 2×3 *pencils.*

nonsquare pencil $A - \lambda B$ is equivalent to one of the following two forms:

$$\begin{bmatrix} 0 & I_m \end{bmatrix} - \lambda \begin{bmatrix} I_m & 0 \end{bmatrix}, (m < n) \quad \text{and} \quad \begin{bmatrix} 0 \\ I_n \end{bmatrix} - \lambda \begin{bmatrix} I_n \\ 0 \end{bmatrix}, (m > n).$$

Square pencils are generically regular, i.e., $\det(A - \lambda B) = 0$ if and only if λ is an eigenvalue. The generic singular pencils of size $n \times n$ have the Kronecker structures [43]:

$$\mathrm{diag}(L_j, L_{n-j-1}^T), \quad j = 0, \ldots, n-1.$$

All generic pencils have full normal rank, i.e., $\mathrm{nrk}(A - \lambda B) = \min(m, n)$. However, a pencil can have full normal rank without being generic. An $m \times n$ pencil with $m < n$ has full normal rank if and only if it has no L_k^T blocks. Similarly, if $m > n$ the pencil has full normal rank if and only if it has no L_k blocks. Finally, a square pencil ($m = n$) has full normal rank if and only if it has no singular blocks.

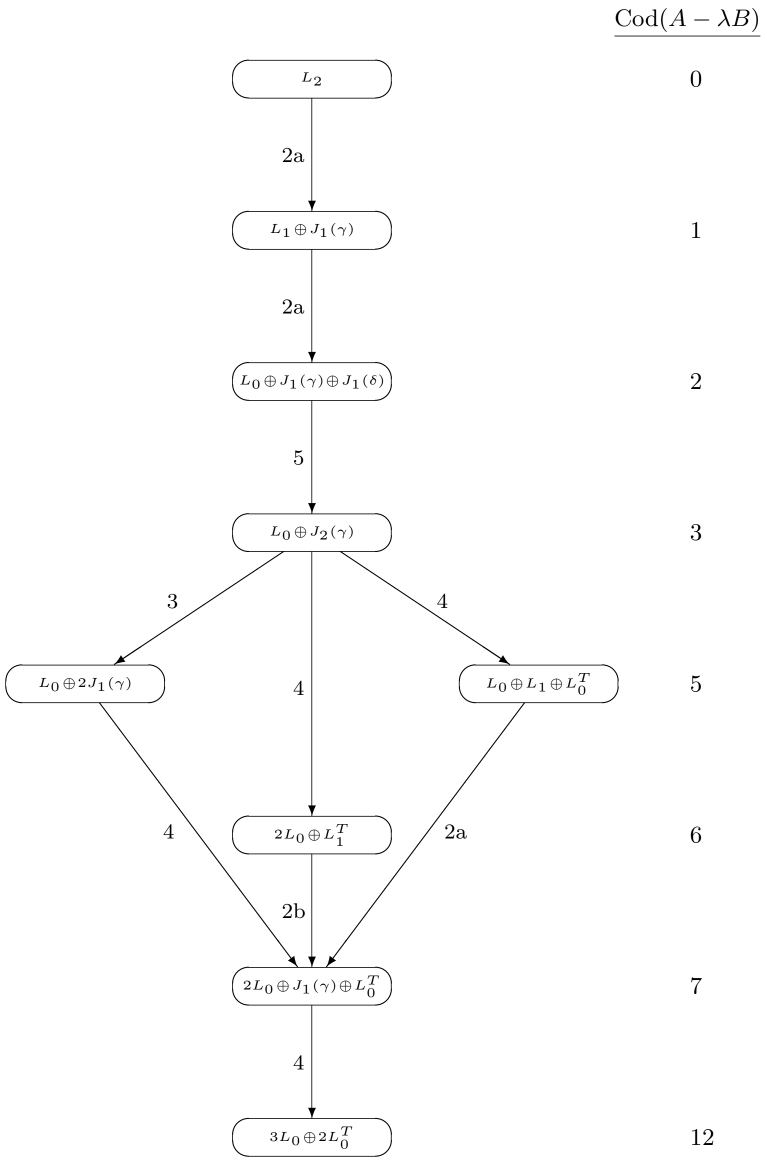

$\mathrm{Cod}(A - \lambda B)$
L_2
0
2a
$L_1 \oplus J_1(\gamma)$
1
2a
$L_0 \oplus J_1(\gamma) \oplus J_1(\delta)$
2
5
$L_0 \oplus J_2(\gamma)$
3
3
4
4
$L_0 \oplus 2J_1(\gamma)$
$L_0 \oplus L_1 \oplus L_0^T$
5
4
$2L_0 \oplus L_1^T$
2a
6
2b
$2L_0 \oplus J_1(\gamma) \oplus L_0^T$
7
4
$3L_0 \oplus 2L_0^T$
12

FIG. 3.4. *Closure hierarchy of bundles for the set of* 2×3 *pencils.*

Next we consider full normal rank pencils with only L_k or L_k^T blocks in their KCF. Let us assume that $m < n$; otherwise we can just perform the same process on the transposed pencil. The $\mathcal{R}$ partition corresponding to the generic pencil is $(r_0, r_1, \cdots, r_\alpha, r_{\alpha+1})$, where

$$r_0 = r_1 = \cdots = r_\alpha = d \quad \text{and } r_{\alpha+1} = m \bmod d$$

for $d = n - m$ and $\alpha = \lfloor m/d \rfloor$. Notice that $r_i = 0$ for $i > \alpha + 1$. Then we have the following corollary of Theorem 3.2.

COROLLARY 3.4. *The dominance ordering of* $\mathcal{R} = (r_0, r_1, \ldots, r_\alpha, r_{\alpha+1})$ *with* $r_0 = d = n - m > 0$ *kept fixed defines the closure hierarchy of the set of* $m \times n$ *matrix pencils with only* L_k *blocks.*

TABLE 3.1
Partitions for deciding closure relations between sample 2×3 pencils.

	$\mathcal{R}$	$\mathcal{L}$	$\mathcal{J}_\gamma$	$\mathcal{J}_\delta$	r_0	$\mathrm{nrk}(\cdot)$
P_1	$(1,1,0,\ldots)$	$(0,\ldots)$	$(1,0,\ldots)$	$(0,\ldots)$	1	2
P_2	$(1,0,\ldots)$	$(0,\ldots)$	$(1,1,0,\ldots)$	$(0,\ldots)$	1	2
P_3	$(1,0,\ldots)$	$(0,\ldots)$	$(1,0,\ldots)$	$(1,0,\ldots)$	1	2
P_4	$(2,0,\ldots)$	$(1,1,0,\ldots)$	$(0,\ldots)$	$(0,\ldots)$	2	1

	$\mathcal{R}+\mathrm{nrk}(\cdot)$	$\mathcal{L}+\mathrm{nrk}(\cdot)$	$\mathcal{J}_\gamma+r_0$	$\mathcal{J}_\delta+r_0$
P_1	$(3,3,2,2,\ldots)$	$(2,2,2,2,\ldots)$	$(2,1,1,1,\ldots)$	$(1,1,1,1,\ldots)$
P_2	$(3,2,2,2,\ldots)$	$(2,2,2,2,\ldots)$	$(2,2,1,1,\ldots)$	$(1,1,1,1,\ldots)$
P_3	$(3,2,2,2,\ldots)$	$(2,2,2,2,\ldots)$	$(2,1,1,1,\ldots)$	$(2,1,1,1,\ldots)$
P_4	$(3,1,1,1,\ldots)$	$(2,2,1,1,\ldots)$	$(2,2,2,2,\ldots)$	$(2,2,2,2,\ldots)$

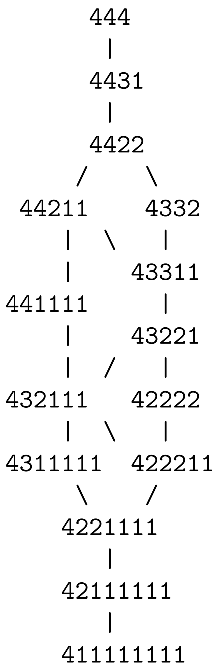


FIG. 3.5. *Covering relationship for $\mathcal{R} = (4,4,4)$ with $r_0 = 4$ kept fixed.*

We illustrate the corollary for the set of 8×12 pencils. The generic Kronecker structure is $4L_2$ ($d = 4, \alpha = 2, m \bmod d = 0$) which gives $\mathcal{R} = (4,4,4)$. The example is chosen so that parts of the dominance ordering for $n = 8$ can be reused (see Figure 2.1). The dominance ordering for $\mathcal{R}$ with r_0 kept fixed is displayed in Figure 3.5. Note the diagram is not symmetric since we are using only a sublattice from Figure 2.1.

A similar result holds for the set of $m \times n$ matrix pencils with a regular part of fixed Jordan structure besides L_j blocks in the KCF. The $\mathcal{R}$ partition corresponding to the most generic pencil with a $k \times k$ regular part is similar to (3.3), where now $d = m - n - k$ and $\alpha = \lfloor (m-k)/d \rfloor$. Since we can write P_1 as $P_1^{(1)} \oplus P_1^{(2)}$, where $P_1^{(1)}$ and $P_1^{(2)}$ correspond to the regular and right singular parts, respectively, we can apply Corollary 3.4 to $P_1^{(2)}$ which defines the $\mathcal{R}$ partition.

Corollary 3.5. *The dominance ordering of* $\mathcal{R} = (r_0, r_1, \ldots, r_\alpha, r_{\alpha+1})$ *with* $r_0 = n - m - k > 0$ *kept fixed defines the closure hierarchy of the set of* $m \times n$ *matrix pencils with a fixed Jordan structure* (*regular part*) *of size* $k \times k$ $(0 \leq k \leq m)$ *and* L_j *blocks only in the KCF.*

These sets of matrix pencils have important applications in linear systems theory. Let $A - \lambda B = [F, C] - \lambda[E, 0]$, where $E\dot{x}(t) = Fx(t) + Cu(t)$ is a linear system with m states and p controls. Then Corollary 3.4 gives the closure hierarchy for the the set of completely controllable systems with m states and p controls. Similarly, Corollary 3.5 gives the closure hierarchy for the sets of linear systems with k uncontrollable modes with fixed Jordan structure.

3.4. Interesting nearness problems. One motivation for our work on versal deformations (see Part I [19]) and stratifications of orbits and bundles (the present paper) was to get an improved understanding of important nearness problems, such as

- closest degenerate (nongeneric) pencil of a generic $A - \lambda B$,
- closest matrix pencil with a specified Kronecker structure,
- closest neighbors (covering pencils) of a given $A - \lambda B$.

Several of these problems have interesting applications in linear system theory. For example, if we add the restriction that the closest degenerate pencil to a generic $m \times n$ pencil (with $m < n$) should have a regular part, then the first problem corresponds to finding the closest uncontrollable system (see also section 3.3).

The closure hierarchy lattice gives one kind of answer to these nearness problems for equivalence orbits of pencils, where we use the codimension instead of the Euclidean distance. We make use of the theorem for covering pencils to prove the following statement.

Theorem 3.6. *Let* $m < n$. *Then the* $m \times n$ *pencils with regular part form a codimension* $n - m + 1$ *stratified submanifold of all pencils equal to the closure of the orbit of*

$$J_1(\gamma) \oplus A_1 - \lambda B_1, \quad \text{where} \quad A_1 - \lambda B_1 = \operatorname{diag}(L_{\bar{\alpha}}, \ldots, L_{\bar{\alpha}}, L_{\bar{\alpha}+1}, \ldots, L_{\bar{\alpha}+1}), \tag{3.4}$$

$\bar{\alpha} = \lfloor (m-1)/d \rfloor$, $d = n - m > 0$, *is the total number of L blocks, and the number of* $L_{\bar{\alpha}+1}$ *blocks is* $(m - 1) \bmod d$. *Therefore, the nearest pencil with a regular part to a generic pencil is generically of the form* (3.4).

Proof. First, we can apply rule (1a) of Theorem 3.2 only until we get a single largest L_k block. Then we apply rule (2a) with the implication that $L_k \oplus \emptyset \rightarrow L_{k-1} \oplus J_1(\gamma)$. The codimension of $J_1(\gamma) \oplus A_1 - \lambda B_1$ is $n - m + 1$. ◻

Notably, $A_1 - \lambda B_1$ is the generic pencil of size $(m-1) \times (n-1)$ and $J_1(\gamma)$ is a Jordan block of size one with an unspecified eigenvalue γ.

We illustrate Theorem 3.6 for 7×12 pencils. In Figure 3.6 we show a sublattice corresponding to equivalence orbits of codimension ≤ 8. We see that $L_0 \oplus 2L_1 \oplus 3L_2$ has the least nonzero codimension (= 3) (Corollary 3.4), and $4L_1 \oplus L_2 \oplus J_1(\gamma)$ has codimension 6 and is the most generic pencil with a regular part (Theorem 3.6). We remark that no structures with a regular part and codimension less than 8 can be found by following the "empty" arc from $L_0 \oplus 2L_1 \oplus L_2 \oplus L_3$, since application of rule (2a) here gives $L_0 \oplus 2L_1 \oplus L_2 \oplus L_3$ and $2L_0 \oplus 2L_2 \oplus L_3$, with codimensions 8 and 9, respectively.

Given a generic $m \times n$ pencil we can apply Theorem 4.2 in our Part I paper [19] to get lower bounds on the distance (measured in the Frobenius norm) to the closest nongeneric pencils of codimension $n - m + 1$.

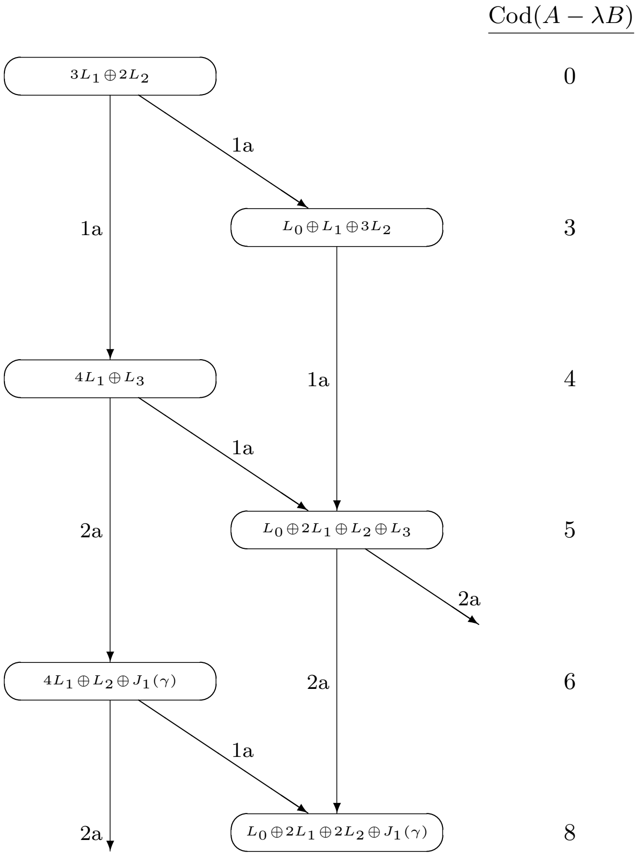

$\mathrm{Cod}(A - \lambda B)$
0
3
4
5
6
8
$3L_1 \oplus 2L_2$
$L_0 \oplus L_1 \oplus 3L_2$
$4L_1 \oplus L_3$
$L_0 \oplus 2L_1 \oplus L_2 \oplus L_3$
$4L_1 \oplus L_2 \oplus J_1(\gamma)$
$L_0 \oplus 2L_1 \oplus 2L_2 \oplus J_1(\gamma)$
1a
1a
1a
1a
2a
2a
2a
1a
2a

FIG. 3.6. *Illustration of Theorem* 3.6. *The closure hierarchy shows that* $4L_1 \oplus L_2 \oplus J_1(\gamma)$ *is the most generic* 7×12 *structure with one eigenvalue.*

3.5. Proofs of Theorems 3.2 and 3.3. Given a pencil P_1, Pokrzywa's Lemma 5 [37] on necessary conditions for covering pencils exhibits a pencil P_2 such that $\overline{\mathcal{O}}(P_1) \supset \overline{\mathcal{O}}(P_2)$, but there may still exist another pencil P such that $\overline{\mathcal{O}}(P_1) \supset \overline{\mathcal{O}}(P) \supset \overline{\mathcal{O}}(P_2)$, i.e., Pokrzywa's rules do not guarantee a cover. The pencils found by his lemma, however, include all pencils P_2 that are covered by P_1; therefore the lemma includes the necessary conditions for covering pencils. We prove Theorem 3.2 by showing that we have included all possible restrictions to the rules without missing any links. For each rule in the proof we denote Pokrzywa's corresponding rule [37, Lem. 5] as (P1), (P2), etc., and consider them in terms of coin moves.

Proof of Theorem 3.2.

(1) (P1) is a rightward coin move in $\mathcal{R}$ (or $\mathcal{L}$) that is consistent with the columns being monotonically ordered. The restriction to a *minimum* rightward coin move precludes the possibility of reaching the same state with another sequence of moves.

(2) (P2) is a coin move from $\mathcal{R}$ (or $\mathcal{L}$) to $\mathcal{J}_{\mu_i}$ for any μ_i. In Theorem 3.2, the reason for moving the *rightmost* coin in $\mathcal{R}$ (or $\mathcal{L}$) is that if we move a coin c that is not the rightmost one, then the same partition can be found by a series of moves (move the coin c using rule (1) until it is in the rightmost

position and then move it to the $\mathcal{J}_{\mu_i}$ partition). This shows that we can generate the same partition with other partitions in between. Similarly, the reason for placing the coin in the *rightmost* position of $\mathcal{J}_{\mu_i}$ is that a partition obtained by placing it in any other position can be obtained by first placing it in the rightmost position and then applying rule (3).

(3) (P3) is a leftward coin move in $\mathcal{J}_{\mu_i}$ that is consistent with the columns being monotonically ordered. Our restriction to a *minimum* leftward coin move is obvious.

(4) (P4) is the removal of a row of coins from one or several eigenvalues. Add one more coin to these $\hat{k}$ coins and distribute all $\hat{k}+1$ coins from left to right in $\mathcal{R}$ and $\mathcal{L}$. There are several restrictions. By picking the *longest row* from each of the $\mathcal{J}_{\mu_i}$, we can always move coins back again (using rule (2)) in order to find the partitions we would have found by removing a shorter row. This is also the reason why we pick the longest row for *all eigenvalues*; if we want to, we can bring them back for all but one eigenvalue. The restriction that each column of $\mathcal{R}$ and $\mathcal{L}$ must have one coin each is required, since otherwise we could obtain the same sequence by first moving more coins to the $\mathcal{J}$ partitions using rule (2) and then applying rule (4).

It is obvious that these four rules are now minimal under these restrictions. □

The proof of Theorem 3.3 is based on the fact that if one orbit is covered by another, then also its bundle is in the closure of the bundle of the other, but the covering relation may be overruled by the possibility of eigenvalues coalescing.

Proof of Theorem 3.3. Rule (5) follows from the matrix case, and since the bundles are unions of orbits, all the covering relations for orbits are also valid for the bundle case, as long as the same operation cannot be performed in more than one step using rule (5). Since it is obvious that the operations corresponding to rules (1) and (3) cannot be done in more steps using rule (5), we only have to focus on rules (2) and (4).

Rule (2) in Theorem 3.2 allows a coin to be moved to any eigenvalue μ_i, but for bundles this can be done in two steps: move the coin to a new eigenvalue μ_j and apply rule (5) on μ_i and μ_j, i.e., append the coin from μ_j to the longest row of coins for μ_i (now $= \mu_j$).

The next question is whether rule (4) can be replaced with a sequence involving rule (5). If there is only one eigenvalue, rule (5) is not applicable, and if each eigenvalue has at least two Jordan blocks, then rule (5) must necessarily decrement the number of distinct eigenvalues while rule (4) does not. Otherwise if an eigenvalue has only one Jordan block, one may apply rule (5) before rule (4) to achieve the same result as a single application of rule (4). □

4. Empowering the staircase algorithm with stratifications.

The staircase algorithm is a powerful tool for computing the Kronecker structure of an $m \times n$ pencil $A - \lambda B$ [4, 8, 32, 35, 34, 41, 44]. The reduction of $A - \lambda B$ into *generalized Schur form* requires several applications of the staircase algorithm. Typically, the first application extracts the right structure and the Jordan structure of the zero eigenvalue using a finite sequence of orthogonal (unitary) equivalence transformations. This decomposition is called the RZ-staircase form (RZ for "right-zero").

In step k $(= 1, 2, \ldots)$ of the first phase, GUPTRI [15, 16] computes the RZ form by determining m_k = dimension of the column nullspace of $A^{(k)}$ and $m_k - s_k$ = dimension of the common column nullspace of $A^{(k)}$ and $B^{(k)}$. Here, $A^{(1)} = A$ and $B^{(1)} = B$ and $(A^{(k)}, B^{(k)})$ for $k > 1$ correspond to the deflated matrix pair obtained after the equivalence transformation in step $k-1$. The structure indices (RZ-indices) display the Kronecker structure as follows:

$$m_k - s_k = \text{ number of } L_{k-1} \text{ blocks}, \quad s_k - m_{k+1} = \text{ number of } J_k(0) \text{ blocks}.$$

The Jordan structure associated with a finite but nonzero eigenvalue is obtained by applying the RZ-staircase algorithm to a shifted pencil. One way to find the left structure is to apply the same algorithm to the transposed pencil. Another way is to directly determine the sizes of the corresponding row nullspaces as done in the GUPTRI algorithm. Then by working on $B - \mu A$ we get the LI-staircase form and now $m_k - s_k$ and $s_k - m_{k+1}$, which are the number of L_{k-1}^T and N_k blocks, respectively, define the LI-indices. Applying the RZ-staircase algorithm to $B - \mu A$ gives the right structure and the Jordan structure of ∞ (RI-indices). Similarly, we can get the left structure and the Jordan structure of zero (LZ-indices) by applying the LI-staircase algorithm to $A - \lambda B$. All combinations (RZ, RI, LI, and LZ) are possible.

Knowing the RZ- and LI-indices we can easily extract the integer sequences (partitions) $\mathcal{R}, \mathcal{L}$, and $\mathcal{J}_{\mu_i}$ or the corresponding staircase indices (R, L, Z, and I). For example, the $\mathcal{R}$ and $\mathcal{J}_0$ partitions are obtained from the RZ-indices as

$$r_{i-1} = \sum_{k=i}^{\infty} m_k - s_k \quad \text{and} \quad j_i = \sum_{k=i}^{\infty} s_k - m_{k+1}.$$

4.1. Modified staircases and covering pencils. In the following we make an algorithmic application of Theorem 3.2. Given a pencil P_1 and the staircase indices defining its Kronecker structure, we want to find *all* pencils P_2 covered by P_1. We can therefore, for example, give the user a selection of choices or perhaps choose one automatically. The four rules in Theorem 3.2 correspond to different structure transitions. Rules (3) and (1) correspond to finding a covering orbit for nilpotent matrices (Corollary 2.3) and full normal rank pencils with only L (or L^T) blocks (Corollary 3.4), respectively. Rule (2) is applicable only if there exists a unique largest L_j (or L_j^T) block in P_1. Then the size of that block is decreased by one, while the size of the largest Jordan block (possibly 0×0) for one eigenvalue is increased by one. Rule (4) replaces the regular structure consisting of the largest Jordan blocks associated with all eigenvalues in P_1 by a generic square singular part. The new L and L^T blocks in P_2 must be at least as large as the corresponding largest singular blocks in P_1.

We assume that the left and right structures are captured only in the structure indices corresponding to one eigenvalue (possibly different for left and right structures). The remaining structure indices capture only Jordan structures, e.g., the RZ-indices associated with an eigenvalue μ_i reduces to Z-indices ($m_k = s_k$).

We propose that a nice user interface based on Algorithm 4.1 should be available to the user.

ALGORITHM 4.1. For all valid coin moves from column j to column k in the appropriate integer sequences ($\mathcal{R}, \mathcal{L}$, and $\mathcal{J}_{\mu_i}$) of rules (1)–(3) in Theorem 3.2, the staircase indices are adjusted as follows. (Remember that $\mathcal{R}$ and $\mathcal{L}$ start counting columns from 0 but $\mathcal{J}_{\mu_i}$ starts from 1.) We use a right arrow (→) to show how one block in the KCF is transferred to another.

(1a) Let m_k and s_k be RZ-indices (or RI-indices). Then $m_{j+1} := m_{j+1} - 1$, $s_j := s_j - 1$ ($L_j \rightarrow L_{j-1}$), $m_{k+1} := m_{k+1} + 1$, and $s_k := s_k + 1$ ($L_{k-1} \rightarrow L_k$).

(1b) Same as item (1a), where now m_k and s_k are LI-indices (or LZ-indices).

(2a) Let m_k and s_k be the RZ-indices (or RI-indices) associated with an eigenvalue $\mu \in \overline{\mathbf{C}}$ (RI-indices if $\mu = \infty$). Then $m_{j+1} := m_{j+1} - 1$, $s_j := s_j - 1$ ($L_j \to L_{j-1}$), $m_k := m_k + 1$, and $s_k := s_k + 1$ ($J_{k-1}(\mu) \to J_k(\mu)$).

(2b) Same as item (2a), where now m_k and s_k are the LZ-indices (or LI-indices) associated with $\mu \in \overline{\mathbf{C}}$.

(3) Let m_k and s_k be any of the staircase indices (RZ, RI, LI, or LZ) associated with $\mu \in \overline{\mathbf{C}}$. Then $m_j := m_j - 1$, $s_j := s_j - 1$ ($J_j(\mu) \to J_{j-1}(\mu)$), $m_k := m_k + 1$, $s_k := s_k + 1$ ($J_{k-1}(\mu) \to J_k(\mu)$).

For all valid coin moves in the appropriate integer sequences ($\mathcal{R}, \mathcal{L}$, and $\mathcal{J}_{\mu_i}$) of rule 4 in Theorem 3.2, the staircase indices are adjusted as follows.

(4) Each valid coin move is defined by k and t in the theorem and the following operations replace the selected $k \times k$ regular part with a generic square singular pencil ($J_{k_1}(\mu_i) \oplus J_{k_2}(\mu_1) \oplus \cdots \oplus J_{k_p}(\mu_p) \to L_t \oplus L_{k-t-1}^T$). Here, k_i is the size of the largest Jordan block of μ_i and $k = \sum k_i$.
   - Repeat for all p eigenvalues μ_i: Let m_k and s_k be the RZ-indices (or RI-indices) of $\mu_i \in \overline{\mathbf{C}}$. Then $s_\iota := s_\iota - 1$, $m_\iota := m_\iota - 1$ for $\iota = 1, \ldots, k_i$.
   - Update RZ-indices (or RI-indices) with respect to new L_t block: $m_\iota := m_\iota + 1$ for $\iota = 1, \ldots, t+1$, $s_\iota := s_\iota + 1$ for $\iota = 1, \ldots, t$ (if $t > 0$).
   - Update LZ-indices (or LI-indices) with respect to new L_{k-t-1}^T block: $m_\iota := m_\iota + 1$ for $\iota = 1, \ldots, k-t$, $s_\iota := s_\iota + 1$ for $\iota = 1, \ldots, k-t-1$ (if $k-t-1 > 0$).

Each valid application of any of the rules (1)–(4) results in a pencil P_2 such that $\mathcal{O}(P_1)$ covers $\mathcal{O}(P_2)$ and each P_2 is on a different branch in the closure lattice. Starting with a generic pencil, repeated applications of the stratification-enhanced algorithm will give us the complete closure hierarchy. Given m_k and s_k corresponding to any of the staircase forms of an arbitrary $m \times n$ pencil up to a certain point, the most generic object is the one where the remaining $\tilde{m} \times \tilde{n}$ pencil is generic. In other words, based on the information obtained up to this point, we know that the pencil is in the closure of the orbit corresponding to this situation. The values of $\tilde{m}$ and $\tilde{n}$ determine the KCF and the staircase indices of the remaining generic pencil (see section 3.3). Any application of the rules (1)–(4) will result in a less generic pencil. Note that the number of different orbits in the closure lattice is exponentially growing as a function of the problem size (m, n), so the algorithm is recursively applied only a few steps if m and n are large.

Similarly, given P_1 it is possible to characterize a pencil P_2 such that $\mathcal{O}(P_2)$ covers $\mathcal{O}(P_1)$. Of course, this will impose different prerequisites on and changes of P_1's structure indices. Moreover, algorithmic applications of Theorem 3.3 for finding covering bundles can be formulated similarly. The algorithmic details are omitted here.

For an illustration of the stratification-enhanced staircase algorithm we return to the examples in Figure 3.2. In Figure 4.1 we display staircase form transitions corresponding to different blocks of P_1 with KCF $L_0 \oplus L_1 \oplus L_2 \oplus J_1(\mu_1) \oplus J_4(\mu_1) \oplus J_3(\mu_2) \oplus L_1^T \oplus L_3^T$. Each of the six cases illustrates the structure index changes imposed by one of the rules of the algorithm. The diagonal blocks in the staircase forms of size $s_k \times m_k$ reveal the Kronecker structures of P_1 and P_2. In order to keep the picture small and clear, we display only the blocks that are directly affected by the transition from P_1 to P_2. These staircase forms correspond to the coin moves illustrated in Figure 3.2. Following the notation from Figure 2.6 the diamond ($\diamondsuit$ for A and $\diamondsuit_\lambda$ for B) is used to denote a matrix entry that the algorithm forces to zero and thereby

Blocks of P_1 $\longrightarrow$ Blocks of P_2

(1a)
$$\left[\begin{array}{cc|cc|c} -\lambda & & 1 & & \\ & \diamondsuit_\lambda & \spadesuit_\lambda & 1 & \\ \hline & & & -\lambda & 1 \end{array}\right] \longrightarrow \left[\begin{array}{c|c|c|c} -\lambda & 1 & & \\ \hline 0 & \spadesuit_\lambda & 1 & \\ \hline & & -\lambda & 1 \end{array}\right]$$

(1b)
$$\left[\begin{array}{c|c|cc} -\lambda & & & \\ \hline 1 & -\lambda & & \\ \hline & 1 & -\lambda & \\ & & \spadesuit & -\lambda \\ \hline & & \diamondsuit & \\ & & & 1 \end{array}\right] \longrightarrow \left[\begin{array}{c|c|c|c} -\lambda & & & \\ \hline 1 & -\lambda & & \\ \hline & 1 & -\lambda & \\ \hline & & \spadesuit & -\lambda \\ \hline & & 0 & \\ & & & 1 \end{array}\right]$$

(2a)
$$\left[\begin{array}{cc|cc|cc|c} -\lambda & & 1 & & & & \\ & -\lambda & & 1 & & & \\ \hline & & & -\lambda & 1 & & \\ & & \diamondsuit_\lambda & & \spadesuit_\lambda & 1 & \\ \hline & & & & & -\lambda & 1 \\ \hline & & & & & & -\lambda \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc|c|c|c} -\lambda & & 1 & & & & \\ & -\lambda & & 1 & & & \\ \hline & & & -\lambda & 1 & & \\ \hline & & 0 & & \spadesuit_\lambda & 1 & \\ \hline & & & & & -\lambda & 1 \\ \hline & & & & & & -\lambda \end{array}\right]$$

(2b)
$$\left[\begin{array}{cc|cc|cc} \spadesuit & -\lambda & & & & \\ \hline \diamondsuit & & -\lambda & & & \\ & 1 & & -\lambda & & \\ \hline & & 1 & & -\lambda & \\ & & & 1 & & -\lambda \\ \hline & & & & 1 & \\ & & & & & 1 \end{array}\right] \longrightarrow \left[\begin{array}{c|c|cc|cc} \spadesuit & -\lambda & & & & \\ \hline 0 & & -\lambda & & & \\ & 1 & & -\lambda & & \\ \hline & & 1 & & -\lambda & \\ & & & 1 & & -\lambda \\ \hline & & & & 1 & \\ & & & & & 1 \end{array}\right]$$

(3)
$$\left[\begin{array}{cc|c|c|c} -\lambda & & 1 & & \\ & -\lambda & & \spadesuit & \\ \hline & & -\lambda & \diamondsuit & \\ \hline & & & -\lambda & 1 \\ \hline & & & & -\lambda \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc|c} -\lambda & & 1 & & \\ & -\lambda & & \spadesuit & \\ \hline & & -\lambda & 0 & \\ & & & -\lambda & 1 \\ \hline & & & & -\lambda \end{array}\right]$$

(4)
$$\left[\begin{array}{c|c|c|c|ccc} -\lambda & 1 & & & & & \\ \hline & -\lambda & 1 & & & & \\ \hline & & \diamondsuit_\lambda & 1 & & & \\ \hline & & & -\lambda & 1 & & \\ \hline & & & & \tilde{\mu}_2-\lambda & 1 & \\ & & & & & \tilde{\mu}_2-\lambda & 1 \\ & & & & & & \tilde{\mu}_2-\lambda \end{array}\right] \longrightarrow \left[\begin{array}{c|c|c|cccc} -\lambda & 1 & & & & & \\ \hline & -\lambda & 1 & & & & \\ \hline & & 0 & 1 & & & \\ & & & -\lambda & 1 & & \\ & & & & \tilde{\mu}_2-\lambda & 1 & \\ & & & & & \tilde{\mu}_2-\lambda & 1 \\ & & & & & & \tilde{\mu}_2-\lambda \end{array}\right]$$

FIG. 4.1. *RZ- and LI-staircase forms of P_1 and P_2 displayed in Figure* 3.2.

changes the computed Kronecker structure from the KCF of P_1 to the KCF of P_2. The spade ($\spadesuit$ for A and $\spadesuit_\lambda$ for B) in P_1 is a nonzero entry that if not existing can be introduced by an equivalence transformation. If not introduced (i.e., the spade does not appear in P_2), then the KCF of P_2 is even less generic, which corresponds to further applications of the stratification-enhanced algorithm.

For cases (1a), (2a), (3), and (4) Figure 4.1 shows the *RZ*-staircase form of P_1 and P_2. Similarly, the *LI*-staircase form is displayed for cases (1b) and (2b). For completeness, we could have included the *LI*-staircase form for case (4) as well.

Rule	Indices	P_1	P_2
(1a)	RZ	k: 1 2 3 4 5 m_k: 5 3 2 1 0 s_k: 4 2 1 1 0	k: 1 2 3 4 5 m_k: 5 2 2 2 0 s_k: 3 2 2 1 0
(1b)	LI	k: 1 2 3 4 5 m_k: 3 3 2 1 0 s_k: 3 2 2 0 0	k: 1 2 3 4 5 6 m_k: 3 2 2 1 1 0 s_k: 2 2 2 1 0 0
(2a)	RZ	See (1a)	k: 1 2 3 4 5 6 m_k: 5 3 1 1 1 0 s_k: 4 1 1 1 1 0
(2b)	LI	See (1b)	k: 1 2 3 4 5 m_k: 3 3 2 1 0 s_k: 3 2 1 1 0
(3)	RZ	See (1a)	k: 1 2 3 4 m_k: 5 4 2 0 s_k: 4 3 1 0
(4)	RZ	See (1a)	k: 1 2 3 4 m_k: 5 3 2 0 s_k: 4 2 0 0
	LI	See (1b)	k: 1 2 3 4 5 6 m_k: 3 3 2 2 1 0 s_k: 3 2 2 1 0 0

FIG. 4.2. *Examples of structure index changes in the stratification-enhanced staircase algorithm.*

However, we see immediately that the last diagonal block of P_2 is an L_4^T block. Here $\tilde{\mu}_2$ corresponds to an eigenvalue $\mu_2 - \mu_1$ of the shifted pencil $A - (\lambda + \mu_1)B$. The sizes $(m_k \times s_k)$ of the diagonal blocks of these staircase forms reveal the changes in the local structure indices that result after applying rules (1)–(4) in the algorithm. For the RZ-staircase forms we start at the top left corner when listing m_k and s_k for $k = 1, \ldots$. Similarly, for the LI-staircase forms we start at the bottom right corner. More interesting are the corresponding changes in the global structure indices (RZ, LI, etc.) for these examples, which are displayed in Figure 4.2. Without loss of generality we have chosen $\mu_1 = 0$ and $\mu_2 = \infty$ in Figure 4.2.

Applying the GUPTRI algorithm in finite precision arithmetic means that all rank decisions for computing the structure indices are made with respect to a user supplied tolerance which reflects the relative accuracy of the data [15, 16]. Assuming a fixed accuracy of the input data it is possible to increase or decrease the tolerance for rank decisions such that a less generic or a more generic pencil, respectively, is computed. Alternatively, given a Kronecker structure computed by the staircase algorithm we can impose a more degenerate Kronecker structure by applying any of the applicable structure index changes. A stratification-enhanced GUPTRI algorithm can deliver an upper bound on the size of the distance from the pencil P_1 we started with to the pencil P_2 we imposed such that $\mathcal{O}(P_1)$ covers $\mathcal{O}(P_2)$. The other way around, we can start with a pencil P_1 and construct a more generic pencil P_2 by adding perturbations (whose sizes are dependent on the rank decision tolerance) such that $\mathcal{O}(P_2)$ covers $\mathcal{O}(P_1)$.

In infinite precision arithmetic we can always go upwards in the closure hierarchy by adding arbitrary small perturbations. This is normally not the case for going downwards in the hierarchy. See [21] for computable normwise bounds of the smallest perturbations for going downward (or upward) in the closure hierarchy of the set of 2×3 pencils.

5. The abstract algebra of matrix pencils.

We give a high-level view of the proofs of Theorem 3.1, mentioning a few new conjectures that we have solved. The Pokrzywa proof uses ordinary linear algebra notation; the algebraic notation by Bongartz would be foreign to many numerical readers. Moreover, we provide a quick introduction to narrow the gap between the algebra and numerical communities. The elegance in the algebraic approach is the unifying treatment obtained for the Jordan, Kronecker, echelon, and many other forms.

5.1. Closure relations, inner products, and codimension counts.

While the covering relationships might be thought of as combinatorial, the closure relations, which are statements about geometry, are derived mainly by algebraic techniques. As discussed earlier, there have been two independent derivations [37, 9] of the closure hierarchy. We suspected that the two very different looking proofs might be somehow "isomorphic," particularly since both count the dimension of the space of solutions to test homogeneous equations.

To be more precise, consider the two *inner products on Kronecker structures* for pencils $P_1 = A_1 - \lambda B_1$ and $P_2 = A_2 - \lambda B_2$,

$$\langle P_1, P_2 \rangle_1 = \dim\{V : A_2 V B_1^T = B_2 V A_1^T\},$$

defined by Pokrzywa [37], and the already defined (before without subscript (3.2))

$$\langle P_1, P_2 \rangle_2 = \dim\{(U, V) : U P_1 = P_2 V\}.$$

The inner product $\langle P_1, P_2 \rangle_2$ is used by Bongartz, who generalizes techniques of Abeasis and Del Fra [1], and by Riedtmann [38], who studied the dimension of the linear space of homomorphisms ($\dim \operatorname{Hom}(P_1, P_2) = \langle P_1, P_2 \rangle_2$ in our case) between path algebra modules (see section 5.2).

Using Kronecker products we can express the inner products as

$$\langle P_1, P_2 \rangle_1 = \dim\{x : T_1 x = 0\} \quad \text{and} \quad \langle P_1, P_2 \rangle_2 = \dim\{y : T_2 y = 0\},$$

where

$$T_1 = \begin{bmatrix} B_1 \otimes A_2 - A_1 \otimes B_2 \end{bmatrix}, \quad x = \operatorname{vec}(V),$$

and

$$T_2 = \begin{bmatrix} A_1^T \otimes I_m & -I_n \otimes A_2 \\ B_1^T \otimes I_m & -I_n \otimes B_2 \end{bmatrix}, \quad x = \begin{bmatrix} \operatorname{vec}(U) \\ \operatorname{vec}(V) \end{bmatrix}.$$

Thus, indeed we have that

$$\langle P_1, P_2 \rangle_1 = \dim \mathcal{N}(T_1) \quad \text{and} \quad \langle P_1, P_2 \rangle_2 = \dim \mathcal{N}(T_2),$$

where $\mathcal{N}(\cdot)$ denotes the nullspace of a matrix. It is clear that with either inner product, if $\overline{\mathcal{O}}(P_1) \supseteq \overline{\mathcal{O}}(P_2)$, then $\langle P_1, T\rangle \leq \langle P_2, T\rangle$ for any test pencil T. Both Pokrzywa and Bongartz prove the converse, giving necessary and sufficient conditions. Furthermore, both observe that one need consider only the indecomposable blocks $T \in \{L_k, L_k^T, J_k\}$ as test pencils giving three sequences of conditions. Explicit formulas may be found in Demmel and Edelman (under the term "interaction") [13] and Beitia and Gracia [5, Thm. 4.5] although these papers did not make the connection to closure relations.

Since the inner product is bilinear, it is sufficient to display a table where the arguments are indecomposable blocks:

$\langle P_1, P_2\rangle_1$	L_k	L_k^T	$J_k(\gamma)$
L_j	$j+k+1$	$(k-j)_+$	k
L_j^T	$(j-k)_+$	0	0
$J_j(\gamma)$	j	0	$\min(j,k)$

$\langle P_1, P_2\rangle_2$	L_k	L_k^T	$J_k(\gamma)$
L_j	$(j+1-k)_+$	0	0
L_j^T	$j+k$	$(k+1-j)_+$	k
$J_j(\gamma)$	j	0	$\min(j,k)$

The inner product on Jordan structures corresponding to different eigenvalues is 0. Here, $J_k(\lambda)$ denotes a Jordan block for a finite or infinite eigenvalue. Notice from the tables that

$$\langle P, L_k\rangle_1 = \langle P^T, L_{k+1}\rangle_2, \quad \langle P, L_k^T\rangle_1 = \langle P^T, L_{k-1}^T\rangle_2, \quad \langle P, J_k\rangle_1 = \langle P^T, J_k\rangle_2.$$

This is not coincidence; we have shown that by eliminating the U from the Bongartz equation involving U and V, one obtains exactly the corresponding Pokrzywa relation, which the reader may notice is symmetric.

Demmel and Edelman were interested in $\operatorname{cod}(P) = \langle P, P\rangle_2 - (m-n)^2$ so as to understand the codimension of $\mathcal{O}(P)$ and the relation to the staircase algorithms for their computation. In our Part I paper [19] we observed that $\langle P, P\rangle_2 = \dim \mathcal{N}(T_2)$. Indeed, when $P = P_1 = P_2$, the tangent space of $P = A - \lambda B$ is the range of the Kronecker product block matrix T_2 [19].

Of course $\operatorname{cod}(P_1) \leq \operatorname{cod}(P_2)$ if $\overline{\mathcal{O}}(P_1) \supseteq \overline{\mathcal{O}}(P_2)$, but the converse does not hold. We at first conjectured that a test pencil approach might work here. The conjecture was that if $\operatorname{cod}(P_1 \oplus T) \leq \operatorname{cod}(P_2 \oplus T)$ for all test pencils T, then $\overline{\mathcal{O}}(P_1) \supseteq \overline{\mathcal{O}}(P_2)$. Unfortunately, this does not hold even for the Jordan case. We found a counterexample consisting of two matrices A_1 and A_2 with Segre characteristics $(5,1,1,1)$ and $(4,3,1)$, respectively. For this example $\operatorname{cod}(A_2 \oplus J_k) \leq \operatorname{cod}(A_1 \oplus J_k)$ for every k, but there is no closure relationship between the orbits of the two matrices (see Figure 2.1).

5.2. Quiver representations and path algebra modules. Perhaps some numerical analysts find it unsatisfying to talk about the Jordan case and then proceed to "analogues" or "generalizations" to the Kronecker case. In fact, the notions of equivalent structures, closure relations, indecomposable blocks, etc., all are elements of an elaborate general theory of *quiver representations* and path algebra modules.

In this theory, the echelon form corresponds to an $\mathbf{A}_2$ quiver with graph consisting of a single arrow ($\bullet \longrightarrow \bullet$), the Jordan form is an $\tilde{\mathbf{A}}_0$ quiver whose graph is a loop ($\circlearrowleft$), and the Kronecker form is an $\tilde{\mathbf{A}}_1$ quiver with two arrows ($\bullet \rightrightarrows \bullet$). A quiver is really a synonym for a directed graph. We obtain a *representation* [18, 39] of the quiver if we associate vertices with vector spaces and arrows with linear maps between the spaces, i.e., matrices. If two vectors are connected tail to tip, then the matrices may be multiplied. A representation of the quiver with three arrows ($\bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet$) is simply three matrices A, B, C that can be multiplied to form CBA. Two representations are said to be equivalent if one can be obtained from the other by a change of basis in the vector space. Thus we have defined similarity of square matrices, strict equivalence of pencils, and many other equivalences with one sentence.

Let E be the incidence matrix of the quiver defined so that E_{ij} is the number of arrows pointing from i to j. The matrix $B = E + E^T$ is independent of the orientation of the arrows. The diagonal elements of B count the number of loops at the node twice. Depending on whether the matrix $2I - B$ is positive definite, semidefinite, or indefinite, the graph is said to be *finite*, *tame*, or *wild*. The finite graphs are known as *Dynkin diagrams* [31]. They correspond to canonical forms built from finitely many blocks, e.g., there are three building blocks for the echelon form: matrices of dimension 1×1, 1×0, and 0×1. The tame quivers include the Jordan and Kronecker forms and are manageable. The wild quivers are more difficult.

We may now define the *path algebra of a quiver*. Formally, it is a vector space generated by elements called paths where multiplication also is defined. A path is simply a sequence of vertices that follow edges. A path of length l may be denoted $(a|\alpha_1, \ldots, \alpha_l|b)$, where a and b are nodes, α_1 is an arrow pointing away from a, the successive arrows point towards each other, and the last arrow points toward b. Paths that connect may be multiplied in the obvious way

$$(a|\alpha_1, \ldots, \alpha_l|b)(b|\beta_1, \ldots, \beta_s|c) = (a|\alpha_1, \ldots, \alpha_l, \beta_1, \ldots, \beta_s|c).$$

Two paths that do not connect are defined to have product 0. Notice that the paths of length 0: $(a|a)$ are idempotent: $(a|a)^2 = (a|a)$ and the sum of all the length 0 paths (one for each node) is the identity.

The Kronecker pencil example is the path algebra for $\tilde{\mathbf{A}}_1$. With the arrows labeled e_1 and e_2, it may be thought of as the four-dimensional space of the form

$$\alpha(1|1) + \beta(1|e_1|2) + \gamma(1|e_2|2) + \delta(2|2)$$

with the path algebra multiplication table:

	$(1\|1)$	$(1\|e_1\|2)$	$(1\|e_2\|2)$	$(2\|2)$
$(1\|1)$	$(1\|1)$	$(1\|e_1\|2)$	$(1\|e_2\|2)$	0
$(1\|e_1\|2)$	0	0	0	$(1\|e_1\|2)$
$(1\|e_2\|2)$	0	0	0	$(1\|e_2\|2)$
$(2\|2)$	0	0	0	$(2\|2)$

We can write such an element as $(\alpha, \beta, \gamma, \delta)$. This algebra is isomorphic to the set of four-dimensional matrices

$$\left(\begin{array}{cc|c} \alpha & 0 & \beta \\ 0 & \alpha & \gamma \\ \hline 0 & 0 & \delta \end{array}\right)$$

with ordinary matrix multiplication. Let A and B be arbitrary $m \times n$ rectangular matrices. We say that vectors v of length $m+n$ form a module over the path algebra of $\tilde{\mathbf{A}}_1$. Define $(\alpha, \beta, \gamma, \delta)v$ to mean

$$\left(\begin{array}{c|c} \alpha I_m & \beta A + \gamma B \\ \hline 0 & \delta I_n \end{array}\right) v. \tag{5.1}$$

It is easy to check that the product

$$(\alpha_1, \beta_1, \gamma_1, \delta_1)(\alpha_2, \beta_2, \gamma_2, \delta_2)v$$

may be computed in either order giving the same answer. (One way requires multiplication in the path algebra; the other is ordinary matrix-vector multiplication.)

There is a one-to-one correspondence between equivalent pencils and modules over the path algebra of $\tilde{\mathbf{A}}_1$, and generally this holds between representations and modules over a path algebra of a quiver. Given two representations of a quiver, if we have linear maps U_i from the first to the second, we say that we have a homomorphism if the diagram composed of the two quivers and the connecting U_i's is commutative, as in the following example:

$$\begin{tikzcd}
\bullet \arrow[r, "A_1"] \arrow[d, "U_1"'] & \bullet \arrow[r, "B_1"] \arrow[d, "U_2"'] & \bullet \arrow[r, "C_1"] \arrow[d, "U_3"'] & \bullet \arrow[d, "U_4"'] \\
\bullet \arrow[r, "A_2"'] & \bullet \arrow[r, "B_2"'] & \bullet \arrow[r, "C_2"'] & \bullet
\end{tikzcd}$$

It is the dimension of the set of homomorphisms between two quivers (dim Hom) that is used explicitly by Bongartz and implicitly by Pokrzywa to obtain the closure relations.

The coin moves also have an algebraic interpretation. Pencils with only L_k blocks correspond to what algebraists call projective modules, while those with only L_k^T blocks are the injective modules. To denote that a matrix A is an extension of A_1 and A_2 (see section 2.3), algebraists write a short exact sequence:

$$0 \to A_1 \to A \to A_2 \to 0.$$

This generalizes to any quiver, and each coin move corresponds to some exact sequence.

This concludes our brief introduction to the algebraic language for these ideas. It is worthwhile to mention that not every equivalence relation in systems theory corresponds to a quiver. The set of matrix pairs (A, B) with A $m \times m$ and B $n \times n$ with the equivalence $(A, B) \sim (U^{-1}AU, U^{-1}B)$ does not correspond to a quiver representation, but if we add the matrix V: $(A, B) \sim (U^{-1}AU, U^{-1}BV)$ then we do have a (wild) quiver [29]. Similarly if we have the matrix quadruples often studied in systems theory [42, 14] with the equivalence relation

$$\begin{pmatrix} P & R \\ O & Q \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} P^{-1} & 0 \\ S & T \end{pmatrix} = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix},$$

we do not have a quiver, but if we omit the matrices R and S, then once again we have a wild quiver.

Acknowledgments. The first author would like to thank Jim Demmel for suggesting this problem one New Year's Eve in Evans Hall (weeks before he met his wife in Evans Hall). We would like to thank Richard Stanley, who always knows the best references and pointers to the literature. We are indebted to Michel Goemans for his help with the proof of Theorem 2.7. Finally we thank Klaus Bongartz for very helpful email discussions, and the minters of Belgian, British, Swedish, French, and American coins.

REFERENCES

[1] S. Abeasis and A. Del Fra, *Degenerations for the representations of a quiver of type $\mathcal{A}_m$*, J. Algebra, 93 (1985), pp. 376–412.

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

[3] V. I. Arnold, S. M. Gusein-Zade, and A. N. Varchenko, *Singularities of Differentiable Maps, Vol.* I, Birkhäuser Boston, Cambridge, MA, 1985.

[4] T. Beelen and P. Van Dooren, *An improved algorithm for the computation of Kronecker's canonical form of a singular pencil*, Linear Algebra Appl., 105 (1988), pp. 9–65.

[5] M. A. Beitia and J.-M. Gracia, *Sylvester matrix equations for matrix pencils*, Linear Algebra Appl., 232 (1996), pp. 155–197.

[6] D. Boley, *Estimating the sensitivity of the algebraic structure of pencils with simple eigenvalue estimates*, SIAM J. Matrix Anal. Appl., 11 (1990), pp. 632–643.

[7] D. Boley, *The Algebraic Structure of Pencils and Block Toeplitz Matrices*, Report TR-96-048, Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN, 1996.

[8] D. Boley and P. Van Dooren, *Placing zeros and the Kronecker canonical form*, Circuits Systems Signal Process., 13 (1994), pp. 783–802.

[9] K. Bongartz, *On degenerations and extensions of finite dimensional modules*, Adv. Math., 121 (1996), pp. 245–287.

[10] F. Chaitin-Chatelin and V. Frayssé, *Lectures on Finite Precision Computations*, SIAM, Philadelphia, PA, 1996.

[11] D. H. Collingwood and W. M. McGovern, *Nilpotent Orbits in Semisimple Lie Algebras*, Van Nostrand Reinhold, New York, 1993.

[12] I. De Hoyos, *Points of continuity of the Kronecker canonical form*, SIAM J. Matrix Anal. Appl., 11 (1990), pp. 278–300.

[13] J. Demmel and A. Edelman, *The dimension of matrices (matrix pencils) with given Jordan (Kronecker) canonical forms*, Linear Algebra Appl., 230 (1995), pp. 61–87.

[14] J. Demmel and B. Kågström, *Accurate solutions of ill-posed problems in control theory*, SIAM J. Matrix. Anal. Appl., 9 (1988), pp. 126–145.

[15] J. Demmel and B. Kågström, *The generalized Schur decomposition of an arbitrary pencil $A-\lambda B$: Robust software with error bounds and applications, part* I*: Theory and algorithms*, ACM Trans. Math. Software, 19 (1993), pp. 160–174.

[16] J. Demmel and B. Kågström, *The generalized Schur decomposition of an arbitrary pencil $A-\lambda B$: Robust software with error bounds and applications, part* II*: Software and applications*, ACM Trans. Math. Software, 19 (1993), pp. 175–201.

[17] H. den Boer and Ph. A Thijsse, *Semi-stability of sums of partial multiplicities under additive perturbation*, Integral Equations Operator Theory, 3 (1980), pp. 23–42.

[18] V. Dlab and C. M. Ringel, *Indecomposable representations of graphs and algebras*, Mem. Amer. Math. Soc., 6 (1996), pp. 155–197.

[19] A. Edelman, E. Elmroth, and B. Kågström, *A geometric approach to perturbation theory of matrices and matrix pencils. Part* I*: Versal deformations*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 653–692.

[20] E. Elmroth, *On the Stratification of the Kronecker Canonical Form*, Report UMINF-95.14, Department of Computing Science, Umeå University, Umeå, Sweden, 1995.

[21] E. Elmroth and B. Kågström, *The set of* 2*-by-*3 *matrix pencils—Kronecker structures and their transitions under perturbations,* SIAM J. Matrix Anal. Appl., 17 (1996), pp. 1–34.

[22] E. R. Gansner, *Acyclic digraphs, Young tableaux and nilpotent matrices*, SIAM J. Alg. Discrete Methods, 2 (1981), pp. 429–440.

[23] F. Gantmacher, *The Theory of Matrices*, *Vols.* I, II, Chelsea, New York, 1959.

[24] M. R. Garey and D. S. Johnson, *Computers and Intractabilty, A Guide to the Theory of NP-Completeness*, Freeman and Company, San Francisco, 1979.

[25] I. Gohberg, P. Lancaster, and L. Rodman, *Invariant Subspaces of Matrices with Applications*, John Wiley, New York, 1986.

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

[27] M. Goresky and R. MacPherson, *Stratified Morse Theory*, Springer-Verlag, Berlin, New York, 1988.

[28] M. Gu, *Finding well-conditioned similarities to block-diagonalize non-symmetric matrices is NP-hard*, J. Complexity, 11 (1995), pp. 377–391.

[29] D. Hinrichsen and D. Prätzel-Wolters, *A wild quiver in linear systems theory*, Linear Algebra Appl., 91 (1987), pp. 143–175.

[30] D. G. Hough, *Explaining and Ameliorating the Ill Condition of Zeros of Polynomials*, Ph.D. thesis, Memo UCB/ERL M77/30, Electronics Research Lab, University of California, Berkeley, CA, 1977.

[31] V. G. Kac, *Montecatini lectures on invariant theory*, in Invariant Theory: Proceedings of the First 1982 session of the Centro Internazionale Matematico Estivo, Springer-Verlag, Berlin, 1983, Lecture Notes in Math., Vol. 996, Proceedings, Montecatini, Italy, 1982.

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

[33] B. Kågström and A. Ruhe, *An algorithm for the numerical computation of the Jordan normal form of a complex matrix*, ACM Trans. Math. Software, 6 (1980), pp. 389–419.

[34] V. B. Khazanov and V. Kublanovskaya, *Spectral problems for matrix pencils, methods and algorithms* I, Sov. J. Numer. Anal. Math. Modelling, 3 (1988), pp. 337–371.

[35] V. Kublanovskaya, *AB-algorithm and its modifications for the spectral problem of linear pencils of matrices*, Numer. Math., 43 (1984), pp. 329–342.

[36] A. S Markus and E. É. Parilis, *The change of the Jordan structure of a matrix under small perturbations*, Linear Algebra Appl., 54 (1983), pp. 139–152.

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

[38] C. Riedtmann, *Degenerations for representations of quivers with relations*, Ann. Sci. École Norm. Sup, 4 (1986), pp. 275–301.

[39] C. M. Ringel, *Tame algebras and integral quadratic forms*, Lecture Notes in Math., 1099, Springer-Verlag, Berlin, 1984.

[40] R. P. Stanley, *Enumerative Combinatorics*, Wadsworth & Brooks/Cole, Monterey, CA, 1986.

[41] P. Van Dooren, *The computation of Kronecker's canonical form of a singular pencil*, Linear Algebra Appl., 27 (1979), pp. 103–141.

[42] P. Van Dooren, *The generalized eigenstructure problem in linear system theory*, IEEE Trans. Automat. Control, AC-26 (1981), pp. 111–129.

[43] W. Waterhouse, *The codimension of singular matrix pairs*, Linear Algebra Appl., 57 (1984), pp. 227–245.

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