On the Stratification of the Kronecker Canonical Form

Erik Elmroth*

May 2, 1995
UMINF-95.14
ISSN-0348-0542

Abstract

The understanding of which Kronecker structures that are close to a given structure is revealed by the Kronecker structure hierarchy, i.e., the stratification of the Kronecker canonical form. For a given matrix pencil $A - \lambda B$, the Kronecker structure hierarchy shows all structures that are within the closure of orbit$(A - \lambda B)$, and each structure, whose orbit's closure contains $A - \lambda B$. In order to gain new insight in the problem of stratification, we give new interpretations of important results by Pokrzywa, for determining closure relations among orbits of Kronecker structures. This is partly done by generalizing classical theorems by Gantmacher. The results are used to derive an algorithm for computation of the complete Kronecker structure hierarchy, or the Kronecker structure hierarchy above or below a given structure. The algorithm is presented in terms of the rank-decisions required in a staircase algorithm, in order to compute the Kronecker structure hierarchy.

Keywords: Kronecker canonical form, Kronecker structure hierarchy, stratification, staircase algorithm, perturbation theory.

---

*Department of Computing Science, Umeå University, S-901 87 Umeå, Sweden. Email address: `elmroth@cs.umu.se` In part supported by the Swedish National Board of Industrial and Technical Development under grant 89-02578P.

1 Introduction

Any algorithm for computation of the Kronecker Canonical form of an m-by-n matrix pencil $A - \lambda B$, computes the exact Kronecker structure of a nearby pencil $A' - \lambda B'$. The distance $\delta \equiv \|(A - A', B - B')\|_E$ is an upper bound on the distance to the closest $(A + \delta A, B + \delta B)$ with the KCF of (A', B'). These algorithms are supposed to compute the most non-generic KCF within distance δ from $A - \lambda B$, and they are said to fail if there is another more non-generic pencil $\hat{A} - \lambda \hat{B}$ within distance δ from the original pencil.

The computation of a KCF can be seen as moving pencils from point to point or manifold to manifold in the $2mn$-dimensional matrix pencil space. During the computation and in the interpretation of the result, it is of great interest to know which structures that are close to a given KCF.

Given $A - \lambda B$, we know mathematically that an arbitrary small perturbation outwards from orbit$(A - \lambda B)$ (the manifold of pencils strictly equivalent to $A - \lambda B$) results in a more generic pencil above $A - \lambda B$ in the closure hierarchy. That is, a structure that is more generic than $A - \lambda B$ and whose orbit's closure includes $A - \lambda B$. By travelling downwards in the closure hierarchy, we find less generic structures, that are in the closure orbit$(A - \lambda B)$.

In order to make the best decisions during the computation of the KCF it is important to understand how the manifolds of different structures relate to each other. Recent contributions to this understanding can be found in [2, 6, 7]. In this contribution we show for the general case how to generate the complete Kronecker structure hierarchy and how to generate the Kronecker structure hierarchy above or below a given KCF.

In Section 2, we briefly review the Kronecker canonical form and the concepts of orbits and their codimension. Section 3 contains a summary of important results by Pokrzywa [11] on the stratification of orbits, i.e., the characterization of closure relations of matrix pencils and necessary conditions on consecutive pencils. In Section 4, we give new interpretations of these results, partly by extending some classical theorems by Gantmacher [8]. The necessary conditions on consecutive pencils are in Section 5 interpreted in terms of rank-decisions of a staircase algorithm for computation of Kronecker structures. In Section 6, our results are put together in an algorithm for computing the complete Kronecker structure hierarchy, or the hierarchy above or below a given structure.

The results in this contribution should be seen as preliminary results for the second part of [6].

2 The Kronecker Canonical Form

Any m-by-n pencil $A - \lambda B$ can by equivalence transformations be transformed into the Kronecker canonical form

$$P^{-1}(A - \lambda B)Q = \mathrm{diag}(L_{\epsilon_1}, \ldots, L_{\epsilon_p}, J_{j_1}(\mu_1), \ldots, J_{j_k}(\mu_k), N_{i_1}, \ldots, N_{i_k}, L^T_{\eta_1}, \ldots, L^T_{\eta_q}),$$

where $J_j(\mu)$ corresponds to a j-by-j Jordan block for the zero or non-zero finite eigenvalue μ and N_j corresponds to a j-by-j Jordan block for the infinite eigenvalue:

$$J_j(\mu) \equiv \begin{bmatrix} \mu-\lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \mu-\lambda \end{bmatrix} \quad \text{and} \quad N_j \equiv \begin{bmatrix} 1 & -\lambda & & \\ & \ddots & \ddots & \\ & & \ddots & -\lambda \\ & & & 1 \end{bmatrix}.$$

The L_j and L_j^T blocks are *singular blocks of right* (column) and *left* (row) *indices of grade j*. These blocks are of size j-by-$(j+1)$ and $(j+1)$-by-j, respectively, and have the form

$$L_j \equiv \begin{bmatrix} -\lambda & 1 & & \\ & \ddots & \ddots & \\ & & -\lambda & 1 \end{bmatrix} \quad \text{and} \quad L_j^T \equiv \begin{bmatrix} -\lambda & & \\ 1 & \ddots & \\ & \ddots & -\lambda \\ & & 1 \end{bmatrix}.$$

The singular blocks have no eigenvalues and there exists a right singular (column) vector that for each λ zeroes out the L_j block identically:

$$\begin{bmatrix} -\lambda & 1 & & \\ & \ddots & \ddots & \\ & & -\lambda & 1 \end{bmatrix} \begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \\ \vdots \\ \lambda^j \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Similarly, there exists a left singular (row) vector that zeroes out the L_j^T block identically:

$$\begin{bmatrix} 1 & \lambda & \lambda^2 & \cdots & \lambda^j \end{bmatrix} \begin{bmatrix} -\lambda & & \\ 1 & \ddots & \\ & \ddots & -\lambda \\ & & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 \end{bmatrix}.$$

If $A-\lambda B$ is regular, the L_j and L_j^T blocks are not present in the Kronecker canonical from.

However, most applications do not require $A-\lambda B$ to be transformed into Kronecker canonical form. Most often it is enough to transfer $A-\lambda B$ to a *generalized Schur form* or similar [1, 3, 4, 9, 10, 12, 14], which reveals the complete Kronecker structure.

If $A-\lambda B$ is m-by-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*). The generic Kronecker structure for $A-\lambda B$ with $d = n - m > 0$ is

$$\operatorname{diag}(L_\alpha, \ldots, L_\alpha, L_{\alpha+1}, \ldots, L_{\alpha+1}), \tag{2.1}$$

where $\alpha = \lfloor m/d \rfloor$, the total number of blocks is d, and the number of $L_{\alpha+1}$ blocks is $m \bmod d$ (which is 0 when d divides m) [12, 2]. The same statement holds for $d = m - n > 0$ if we replace $L_\alpha, L_{\alpha+1}$ in (2.1) by $L_\alpha^T, L_{\alpha+1}^T$. 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-by-n have the Kronecker structures [13]:

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

We define an *orbit* to be the set of *strictly equivalent* pencils in $2mn$-dimensional space:

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

That is, an orbit defines a manifold of pencils with identical Kronecker structures. The dimension of $\mathrm{orbit}(A - \lambda B)$ is equal to the dimension of the tangent space, $\tan(A - \lambda B)$, to the orbit of $A - \lambda B$. The tangent space is defined as

$$f(X, Y) = X(A - \lambda B) - (A - \lambda B)Y, \tag{2.2}$$

where X is an $m \times m$ matrix and Y is an $n \times n$ matrix [2]. Since (2.2) maps a space of dimension $m^2 + n^2$ linearly into a space of dimension $2mn$, the dimension of the tangent space is m^2+n^2-d, where d is the number of (linearly) independent solutions of $f(X, Y) = 0$.

The codimension is the dimension of the space complementary to the tangent space, i.e.,

$$\mathrm{cod}(A - \lambda B) = 2mn - \dim(\tan(A - \lambda B)) = d - (m - n)^2.$$

As shown in [6], the codimension of $A - \lambda B$ can be computed as the number of zero singular values of

$$T \equiv \begin{bmatrix} A^T \otimes I_m & -I_n \otimes A \\ B^T \otimes I_m & -I_n \otimes B \end{bmatrix}.$$

We may also define the normal space, $\mathrm{nor}(A - \lambda B)$, as the space perpendicular to $\tan(A - \lambda B)$. Then $\mathrm{cod}(A - \lambda B)$ is the dimension of $\mathrm{nor}(A - \lambda B)$.

Since the codimension of an orbit depends only on its Kronecker structure, $\mathrm{cod}(A - \lambda B)$ can also be computed by summing the contributions to the codimension from different blocks in the KCF [2].

3 Stratification of Orbits

Given two $m \times n$ matrix pencils $\mathcal{P}_1 = A_1 - \lambda B_1$ and $\mathcal{P}_2 = A_2 - \lambda B_2$ we are interested to know when the closure of $\mathrm{orbit}(\mathcal{P}_1)$ includes the closure of $\mathrm{orbit}(\mathcal{P}_2)$, i.e., $\overline{\mathrm{orbit}}(\mathcal{P}_1) \supset \overline{\mathrm{orbit}}(\mathcal{P}_2)$. As in the matrix case we define $\mathcal{P}_1$ and $\mathcal{P}_2$ to be *consecutive* if $\overline{\mathrm{orbit}}(\mathcal{P}_1) \supset \overline{\mathrm{orbit}}(\mathcal{P}_2)$ and there is no pencil $\mathcal{P}$ such that $\overline{\mathrm{orbit}}(\mathcal{P}_1) \supset \overline{\mathrm{orbit}}(\mathcal{P}) \supset \overline{\mathrm{orbit}}(\mathcal{P}_2)$.

Since the KCF may include both Jordan blocks associated with finite as well as infinite eigenvalues and Kronecker blocks corresponding to left and right minimal indices, the task to get a complete characterization of all possible $m \times n$ Kronecker forms is much more intricate than for the matrix case, where we only have to consider Jordan blocks.

3.1 Characterization of the closure of $\text{orbit}(\mathbf{A}-\lambda\mathbf{B})$

An important contribution to the understanding of the stratification for the Kronecker canonical form was presented by Pokrzywa [11] in 1986. We start by reviewing some of his results (using our notation).

The following theorem gives a characterization of $\overline{\text{orbit}}(\mathcal{P}_1) \supset \overline{\text{orbit}}(\mathcal{P}_2)$ in terms of their Kronecker structures. For a given pencil $\mathcal{P} = A - \lambda B$ let $r_k(\mathcal{P})$ and $l_k(\mathcal{P})$ be the number of L_k (right singular) and L_k^T (left singular) blocks, respectively, and let $r(\mathcal{P})$ be the total number of right singular blocks in the KCF of $\mathcal{P}$. Moreover, let $j_k(\mathcal{P},\mu)$ be the number of $k \times k$ Jordan blocks associated with the eigenvalue μ of $\mathcal{P}$.

Theorem 3.1 *[11] A pencil $\mathcal{P}_2$ is in the closure of* $\text{orbit}(\mathcal{P}_1)$ *if and only if the following inequalities hold:*

$$\sum_k (i-k)_+ \cdot r_k(\mathcal{P}_1) \le \sum_k (i-k)_+ \cdot r_k(\mathcal{P}_2), \tag{3.1}$$

$$\sum_k (i-k)_+ \cdot l_k(\mathcal{P}_1) \le \sum_k (i-k)_+ \cdot l_k(\mathcal{P}_2), \tag{3.2}$$

$$i \cdot r(\mathcal{P}_1) + \sum_k \min(i,k) \cdot j_k(\mathcal{P}_1,\mu) \le i \cdot r(\mathcal{P}_2) + \sum_k \min(i,k) \cdot j_k(\mathcal{P}_2,\mu), \tag{3.3}$$

for all $i = 1, 2, \ldots$ and $\mu \in \overline{\mathbf{C}}$, where $\overline{\mathbf{C}}$ is the complex plane including the point at infinity and $(x)_+ = \max(0,x)$.

Theorem 3.1 gives us necessary and sufficient conditions to decide whether one pencil is in the closure of another. However, to get a complete understanding of the closure hierarchy we also need to be able to decide whether two pencils are consecutive or not, or given a pencil $\mathcal{P}_1$ we want to know how to construct all possible consecutive pencils (above and below in the closure hierarchy).

An instrument for this purpose is the following characterization of two consecutive pencils [11]. If $\overline{\text{orbit}}(\mathcal{P}_1) \supset \overline{\text{orbit}}(\mathcal{P}_2)$ and $\mathcal{P}_1$ and $\mathcal{P}_2$ are consecutive pencils, then $\mathcal{P}_1 = P_1^{(1)} \oplus P_1^{(2)}$ and $\mathcal{P}_2 = P_2^{(1)} \oplus P_2^{(2)}$, where $P_1^{(1)}$ and $P_2^{(1)}$ are equivalent, and $P_1^{(2)}$ and $P_2^{(2)}$ are equivalent to one of the ordered pairs of pencils $(\overline{\text{orbit}}(P_1^{(2)}) \supset \overline{\text{orbit}}(P_2^{(2)}))$ of the following lemma.

Lemma 3.1 *[11] The following closure relations hold:*

(i) $\overline{\text{orbit}}(L_j \oplus L_k) \supset \overline{\text{orbit}}(L_{j-1} \oplus L_{k+1}),\ 1 \le j \le k.$

(ii) $\overline{\text{orbit}}(L_j^T \oplus L_k^T) \supset \overline{\text{orbit}}(L_{j-1}^T \oplus L_{k+1}^T),\ 1 \le j \le k.$

(iii) $\overline{\text{orbit}}(L_{j+1} \oplus J_k(\mu)) \supset \overline{\text{orbit}}(L_j \oplus J_{k+1}(\mu)),\ j,k = 0,1,2,\ldots$ and $\mu \in \overline{\mathbf{C}}$.

(iv) $\overline{\text{orbit}}(L_{j+1}^T \oplus J_k(\mu)) \supset \overline{\text{orbit}}(L_j^T \oplus J_{k+1}(\mu)),\ j,k = 0,1,2,\ldots$ and $\mu \in \overline{\mathbf{C}}$.

(v) $\overline{\text{orbit}}(J_{j-1}(\mu)) \oplus J_{k+1}(\mu)) \supset \overline{\text{orbit}}(J_j(\mu)) \oplus J_k(\mu)),\ 1 \le j \le k$ and $\mu \in \overline{\mathbf{C}}$.

(vi) $\overline{\text{orbit}}(\bigoplus_{i=1}^t J_{k_i}(\lambda_i)) \supset \overline{\text{orbit}}(L_p \oplus L_q^T)$, if $\lambda_i \ne \lambda_j$ for $i \ne j$, $\lambda_i \in \overline{\mathbf{C}}$ and $p + q + 1 = \sum_{i=1}^t k_i$.

Given a pencil $\mathcal{P}_1$ and its KCF we can use Lemma 3.1 to find a pencil $\mathcal{P}_2$ in $\overline{\text{orbit}}(\mathcal{P}_1)$ (i.e., $\mathcal{P}_2$ is below in the hierarchy) such that $\mathcal{P}_1$ and $\mathcal{P}_2$ fulfill necessary conditions for two consecutive pencils. Similarly, using Lemma 3.1 we can find another pencil $\mathcal{P}_2$ such that $\mathcal{P}_1$ is in $\overline{\text{orbit}}(\mathcal{P}_2)$ (i.e., $\mathcal{P}_2$ is above in the hierarchy). In both cases $\mathcal{P}_1$ and $\mathcal{P}_2$ fulfill necessary but not sufficient conditions for two consecutive pencils. Notice that it is allowed to have $J_0(\mu)$ blocks in the left hand side of items (iii), (iv), and (v) [11]. For example, in item (iii) this means that $\overline{\text{orbit}}(L_{j+1}) \supset \overline{\text{orbit}}(L_j \oplus J_1(\mu))$.

Recently, Elmroth and Kågström derived the closure hierarchy (or stratification) of the set of 2×3 pencils [7]. By considering zero, non-zero and infinite eigenvalues separately (which is motivated from an algorithmic point of view) there are 20 structurally different Kronecker structures. From a topological point of view it is natural to only consider different eigenvalues without any special attention to zero and infinite eigenvalues, resulting in only 9 structurally different Kronecker structures. The closure hierarchy graph corresponding to these is shown in Figure 1.

Following [7] we display closure graphs such that orbits (nodes) with the same codimension are displayed on the same horizontal level. The generic case (L_2) is at the highest level and the most non-generic pencil ($3L_0 \oplus 2L_0^T$ which is the 2×3 zero pencil) is at the lowest level. Moreover, along each arc in Figure 1 we display the item in Lemma 3.1 for the necessary condition for consecutive pencils that is fulfilled. From this closure hierarchy it is possible to find examples where the necessary conditions of Lemma 3.1 are not sufficient conditions for consecutive pencils. For example, $\mathcal{P}_1 = L_1 \oplus J_1(\gamma)$ and $\mathcal{P}_2 = L_0 \oplus 2J_1(\gamma)$ fulfill the necessary condition (iii), $\mathcal{P}_1 = L_1 \oplus J_1(\gamma)$ and $\mathcal{P}_2 = L_0 \oplus L_1 \oplus L_0^T$ fulfill the necessary condition (vi), as $\mathcal{P}_1 = L_0 \oplus J_1(\gamma) \oplus J_1(\delta)$ and $\mathcal{P}_2 = 2L_0 \oplus J_1(\gamma) \oplus L_0^T$ do, while in all three cases there exists a pencil $\mathcal{P}$ such that $\overline{\text{orbit}}(\mathcal{P}_1) \supset \overline{\text{orbit}}(\mathcal{P}) \supset \overline{\text{orbit}}(\mathcal{P}_2)$.

4 Characterizing Closure Relations Using Gantmacher Matrices

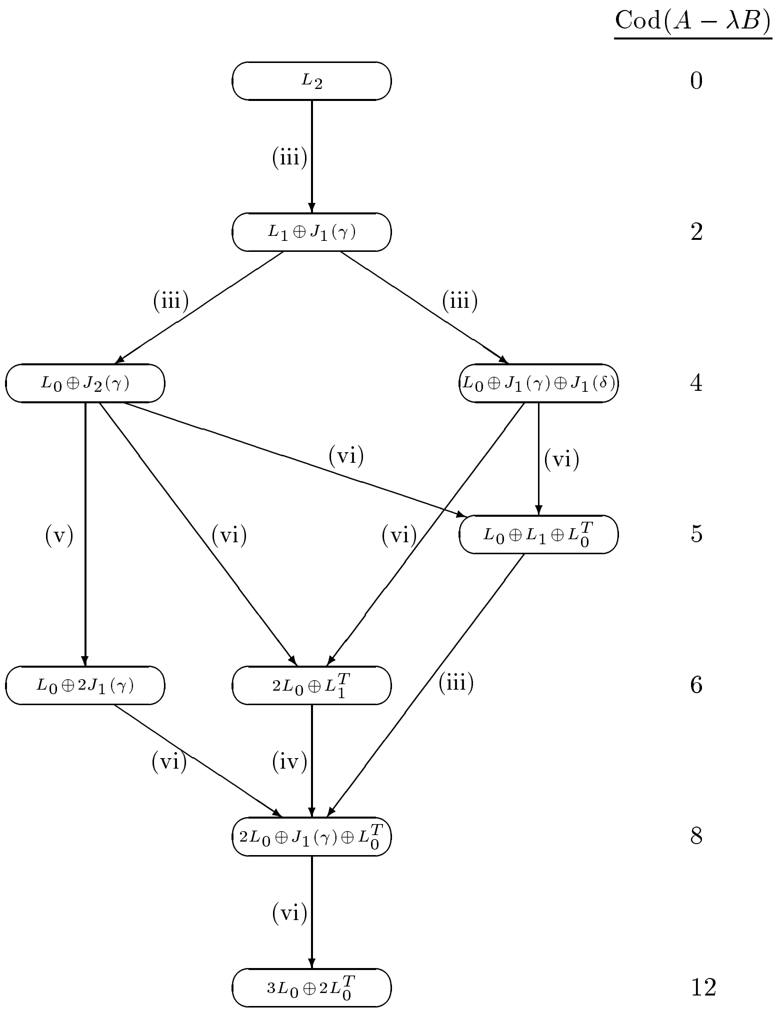
Let the Gantmacher matrix $R[A, B, i]$ of size $(i+2)m \times (i+1)n$ be defined by

$$R[A,B,i] = \begin{bmatrix} A & 0 & \cdots & 0 \\ B & A & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & B & A \\ 0 & \cdots & 0 & B \end{bmatrix},$$

where A and B are $m \times n$ matrices. When it is clear from context we use the abbreviated notation $R[i]$, or $R[\mathcal{P}, i]$ where $\mathcal{P} = A - \lambda B$, for $R[A, B, i]$.

Gantmacher showed that if $R[i-1]$ has full column rank but $R[i]$ does not, then there exists an L_i block in the Kronecker structure of $A - \lambda B$ [8]. Here we extend this result in theorems 4.1–4.2 and we use these extensions in Theorem 4.3 to characterize closure relations in terms of Gantmacher matrices. The objective is to gain more insight in the problem of stratification of orbits of Kronecker structures.

Figure 1: Closure hierarchy graph for the set of 2-by-3 pencils.



Theorem 4.1 *If $R[A,B,i-1]$ has full column rank and $R[A,B,i]$ has a k-dimensional column nullspace, then there exist k L_i blocks in the KCF of $A-\lambda B$.*

Proof. Since $R[i-1]$ has full column rank but $R[i]$ does not (since it has a k-dimensional column nullspace), we know from Gantmacher's results that $A-\lambda B$ is equivalent to the pencil

$$\left[\begin{array}{cc} L_i & 0 \\ 0 & A'-\lambda B' \end{array}\right]. \tag{4.1}$$

By rewriting $R[i]$ using this equivalent pencil and the notation $L_i = L_i^A - \lambda L_i^B$ we have,

$$R[A,B,i] = \left[\begin{array}{cc|cc|cc|cc} L_i^A & 0 & & & & & & \\ 0 & A' & & & & & & \\ \hline L_i^B & 0 & L_i^A & & & & & \\ 0 & B' & 0 & A' & & & & \\ \hline & & \ddots & & \ddots & & & \\ & & & \ddots & & \ddots & & \\ \hline & & & & L_i^B & 0 & L_i^A & 0 \\ & & & & 0 & B' & 0 & A' \\ \hline & & & & & & L_i^B & 0 \\ & & & & & & 0 & B' \end{array}\right], \tag{4.2}$$

where the empty blocks denote zeros. By using

$$U = \left[\begin{array}{cc|cc|c|cc} I_i & 0 & & & & & \\ \hline & & I_i & 0 & & & \\ \hline & & & & \ddots & & \\ \hline & & & & & I_i & 0 \\ \hline 0 & I_{m-i} & & & & & \\ \hline & & 0 & I_{m-i} & & & \\ \hline & & & & \ddots & & \\ \hline & & & & & 0 & I_{m-i} \end{array}\right], \tag{4.3}$$

and

$$V = \left[\begin{array}{c|c|c|c|c|c|c|c} \begin{matrix} I_{i+1} \\ 0 \end{matrix} & & & & \begin{matrix} 0 \\ I_{n-i-1} \end{matrix} & & & \\ \hline & \begin{matrix} I_{i+1} \\ 0 \end{matrix} & & & & \begin{matrix} 0 \\ I_{n-i-1} \end{matrix} & & \\ \hline & & \ddots & & & & \ddots & \\ \hline & & & \begin{matrix} I_{i+1} \\ 0 \end{matrix} & & & & \begin{matrix} 0 \\ I_{n-i-1} \end{matrix} \end{array}\right], \tag{4.4}$$

for row and column permutations, respectively, we have

$$U\,R[A,B,i]\,V = \left[\begin{array}{cc} R[L_i^A, L_i^B, i] & 0 \\ 0 & R[A',B',i] \end{array}\right]. \tag{4.5}$$

It is easily shown that $R[L_i^A, L_i^B, i]$ has full row rank and since it is of dimension $(i^2 + 2i) \times (i^2 + 2i + 1)$, the matrix $R[L_i^A, L_i^B, i]$ has a 1-dimensional column nullspace. Since $R[A, B, i]$ has a k-dimensional column nullspace, the matrix $R[A', B', i]$ must have a $(k-1)$-dimensional column nullspace. If $k - 1 \geq 1$, then there is at least one L_i block in the KCF of $A' - \lambda B'$. By repeating the procedure above, we are in total able to extract k L_i blocks from $A - \lambda B$. $\square$

Theorem 4.2 *If $R[A, B, i]$ has a k-dimensional column nullspace given by k L_i blocks in the KCF of $A - \lambda B$, then these k L_i blocks induce a $k(p-i+1)$-dimensional column nullspace in $R[A, B, p]$ for $p \geq i$.*

Proof. Let $R[A, B, i]$ have a k-dimensional column nullspace given by k L_i blocks. The total nullspace may of course be larger due to existence of other right singular blocks. By using similar reorganizations as in (4.1)–(4.5) of the matrix $R[A, B, i]$, we can obtain a reorganized matrix with k $R[L_i^A, L_i^B, i]$ matrices in its block structure. Similarly, a reorganized $R[A, B, p]$ has k $R[L_i^A, L_i^B, p]$ matrices in its block structure. Since each $R[L_i^A, L_i^B, p]$ matrix has full row rank and dimensions $(p + 2)i \times (p + 1)(i + 1)$, each $R[L_i^A, L_i^B, p]$ has a $(p - i + 1)$-dimensional column nullspace for $p \geq i$. It follows that the k matrices $R[L_i^A, L_i^B, i]$ in total give a $k(p - i + 1)$-dimensional column nullspace in $R[A, B, p]$ for $p \geq i$. $\square$

From Theorem 3.1 we have that the right singular structure of a pencil $\mathcal{P} = A - \lambda B$ is characterized by the expression

$$\sum_k (i - k)_+ \cdot r_k(\mathcal{P}), \text{ for } i = 1, 2, \ldots, \tag{4.6}$$

where $r_k(\mathcal{P})$ is the number of L_k blocks in the KCF of $\mathcal{P}$. For a given value i the sum (4.6) is equal to

$$i \cdot r_0 + (i - 1) \cdot r_1 + \ldots + 1 \cdot r_{i-1}.$$

From theorems 4.1–4.2 we know that $(i - j) \cdot r_j$ is the dimension of the part of the column nullspace of $R[A, B, i - 1]$ that is induced by r_j L_j blocks. It follows that we get an explicit expression for $\mathrm{n}_c(R[i])$, the dimension of the column nullspace of $R[i]$.

Corollary 4.1 *Let $r_k(\mathcal{P})$ be the number of right singular blocks in $\mathcal{P} = A - \lambda B$. Then*

$$\sum_k (i - k)_+ \cdot r_k(\mathcal{P}) \equiv \mathrm{n}_c(R[i - 1]), \tag{4.7}$$

for $i = 1, 2, \ldots$.

To identify the existence of L_i^T blocks in the KCF, Gantmacher stated a dual form of his theorem. It says that if the $(i + 1)m \times (i + 2)n$ matrix

$$L[A, B, i] = \begin{bmatrix} A & B & 0 & \cdots & 0 \\ 0 & A & B & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A & B \end{bmatrix},$$

does not have full row rank but $L[A, B, i-1]$ does, then there exists an L_i^T block in the KCF of $A - \lambda B$.

As before, we also use the abbreviated notations $L[i]$ or $L[\mathcal{P}, i]$ for $L[A, B, i]$. It is straightforward to formulate dual versions of theorems 4.1 and 4.2, resulting in similar explicit expressions for $\mathrm{n}_r(L[i])$, the dimension of the row nullspace of $L[i]$.

Corollary 4.2 *Let $l_k(\mathcal{P})$ be the number of left singular blocks in $\mathcal{P} = A - \lambda B$. Then*

$$\sum_k (i-k)_+ \cdot l_k(\mathcal{P}) \equiv \mathrm{n}_r(L[i-1]), \tag{4.8}$$

for $i = 1, 2, \ldots$.

We have now showed how to interpret the first two inequalities in Theorem 3.1, using Gantmacher matrices. Our interpretation of the third inequality is partly based on some final remarks by Pokrzywa [11]. He notes that if we consider the n-by-n pencil $A - \mu I$, then

$$\sum_k \min(i, k) \cdot j_k(\mathcal{P}, \mu) = \mathrm{n}_c((A - \mu I)^i),$$

which we recognize as the sum of the i first elements in the Weyr characteristics corresponding to the eigenvalue μ [5]. (Since A is square, we also have $\mathrm{n}_c((A - \mu I)^i) = \mathrm{n}_r((A - \mu I)^i)$). For general $A - \lambda B$, let w_j for the eigenvalue μ denote the number of $J_k(\mu)$ blocks for $k \geq j$. Then

$$\sum_k \min(i, k) \cdot j_k(\mathcal{P}, \mu) = \sum_{j=1}^{i} w_j.$$

We are now able to rewrite Theorem 3.1 in terms of Weyr characteristics and Gantmacher matrices.

Theorem 4.3 *A pencil $\mathcal{P}_2$ is in the closure of* $\mathrm{orbit}(\mathcal{P}_1)$ *if and only if the following inequalities hold:*

$$\begin{aligned} \mathrm{n}_c(R[\mathcal{P}_1, i]) &\leq \mathrm{n}_c(R[\mathcal{P}_2, i]), & (4.9)\\ \mathrm{n}_r(L[\mathcal{P}_1, i]) &\leq \mathrm{n}_r(L[\mathcal{P}_2, i]), & (4.10)\\ i \cdot r(\mathcal{P}_1) + \sum_{j=1}^{i} w_j^{(1)} &\leq i \cdot r(\mathcal{P}_2) + \sum_{j=1}^{i} w_j^{(2)}. & (4.11) \end{aligned}$$

for all $i = 0, 1, 2, \ldots$ and $w_j^{(1)}$ and $w_j^{(2)}$, $j = 1, 2, \ldots$, denote the Weyr characteristics corresponding to the eigenvalue $\mu \in \overline{\mathbf{C}}$, for $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively.

Note that we here start form $i = 0$ in order to obtain $R[i]$ and $L[i]$ in (4.10) and (4.10) instead of $R[i-1]$ and $L[i-1]$, as in (4.7) and (4.8). The inequality (4.11) is not affected by this change, since $i = 0$ makes both the left hand side and the right hand side equal to zero.

One immediate observation is that since $w_i = 0$ for i large enough, $r(\mathcal{P}_1) \leq r(\mathcal{P}_2)$ must hold in order to satisfy the third inequality for all i. Indeed, by examining the first two inequalities using Theorem 4.2, we see that they require $\mathcal{P}_2$ to have at least as many L_k and L_k^T blocks as $\mathcal{P}_1$.

5 Characterizing Consecutive Pencils Using the Staircase Algorithm

Using the staircase algorithm (in infinite precision arithmetic) it is possible to compute the (exact) Kronecker structure of a given $m \times n$ pencil. One phase of the algorithm extracts, for example, the Jordan structure of the zero eigenvalue and the right singular structure of $A - \lambda B$ using a finite sequence of orthogonal (unitary) equivalence transformations. In step k $(= 0, 1, \ldots)$ of the first phase, the GUPTRI algorithm [3, 4] (one variant of the staircase algorithm) determines $m_k =$ 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, $A^{(0)} = A$ and $B^{(0)} = B$ and $(A^{(k)}, B^{(k)})$ for $k \geq 1$ correspond to the deflated matrix pair obtained after the equivalence transformation in step $k - 1$. The structure indices display the Kronecker structure as follows:

- $m_k - s_k =$ number of L_k blocks.
- $s_k - m_{k+1} =$ number of $J_{k+1}(0)$ blocks.

Applying the same algorithm to $B - \mu A$ results in the Jordan structure of the infinite eigenvalue and the right singular structure. The Jordan structure (and structure indices) associated with a finite but non-zero eigenvalue is obtained by applying the algorithm to a shifted pencil. One way to find the left singular 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, resulting in the Jordan structure of the infinite eigenvalue and the left singular structure. Then $m_k - s_k$ is the number of L_k^T blocks and $s_k - m_{k+1}$ equals the number of $N_{k+1} \equiv J_{k+1}(\infty)$ blocks.

In Corollary 5.1, we formulate Lemma 3.1 in terms of the structure indices computed by the staircase algorithm. Given a pencil $\mathcal{P}_1$, we show prerequisites on its structure indices for each item in Lemma 3.1 and how they must change in order to characterize a pencil $\mathcal{P}_2$ such that $\overline{\text{orbit}}(\mathcal{P}_1) \supset \overline{\text{orbit}}(\mathcal{P}_2)$ and $\mathcal{P}_1$ and $\mathcal{P}_2$ fulfill necessary conditions for two consecutive pencils. We use an arrow $(\rightarrow)$ to show how one block in the KCF is transferred to another.

Corollary 5.1 *Items* (i)–(vi) *show prerequisites on the m_i and s_i indices of a pencil $\mathcal{P}_1$ and the changes required in these indices to find a pencil $\mathcal{P}_2$, such that $\overline{\text{orbit}}(\mathcal{P}_1) \supset \overline{\text{orbit}}(\mathcal{P}_2)$ and $\mathcal{P}_1$ and $\mathcal{P}_2$ fulfill necessary conditions for two consecutive pencils.*

(i) *For $1 \leq j \leq k$ there exist at least one L_j and L_k in the KCF of $\mathcal{P}_1$ if either $m_j > s_j$, $m_k > s_k$ and $j < k$, or $m_j > s_j + 1$ and $j = k$. $\mathcal{P}_2$ is characterized by the changes $m_j := m_j - 1$, $s_{j-1} := s_{j-1} - 1$ $(L_j \rightarrow L_{j-1})$, $m_{k+1} := m_{k+1} + 1$, and $s_k := s_k + 1$ $(L_k \rightarrow L_{k+1})$.*

(ii) *Similar to item* (i) *but with the m_k and s_k indices replaced by the indices representing the corresponding left singular structure.*

(iii) *For $j, k \geq 0$ there exist at least one L_{j+1} and $J_k(\mu)$ in the KCF of $\mathcal{P}_1$ if $m_{j+1} > s_{j+1}$ and $s_{k-1} > m_k$. The last relation is not applicable if $k = 0$. $\mathcal{P}_2$ is characterized by the changes $m_{j+1} := m_{j+1} - 1$, $s_j := s_j - 1$ $(L_{j+1} \rightarrow L_j)$, $m_k := m_k + 1$, and $s_k := s_k + 1$ $(J_k(\mu) \rightarrow J_{k+1}(\mu))$.*

(iv) *Similar to item* (iii) *but with the* m_k *and* s_k *indices replaced by the indices representing the corresponding left singular structure and Jordan structure.*

(v) *For* $1 \leq j \leq k$ *there exist at least one* $J_{j-1}(\mu)$ *and* $J_{k+1}(\mu)$ *in the KCF of* $\mathcal{P}_1$ *if* $s_{j-2} > m_{j-1}$, $s_k > m_{k+1}$. *The first relation is not applicable if* $j = 1$. $\mathcal{P}_2$ *is characterized by the changes* $m_{j-1} := m_{j-1} + 1$, $s_{j-1} := s_{j-1} + 1$ $(J_{j-1}(\mu) \rightarrow J_j(\mu))$, $m_k := m_k - 1$, *and* $s_k := s_k - 1$ $(J_{k+1}(\mu) \rightarrow J_k(\mu))$.

(vi) *There exist* t *Jordan blocks* $J_{k_i}(\mu_i), (k_i \geq 1)$ *where each Jordan block corresponds to a different eigenvalue* μ_i. *The structure indices for each eigenvalue must fulfill* $s_{k_i - 1} > m_{k_i}$. $\mathcal{P}_2$ *is characterized by the changes* $s_{\delta-1} := s_{\delta-1} - 1$, $m_\delta := m_\delta - 1$, *for* $\delta = 1, \ldots, k_i$ *for the indices corresponding to each of the eigenvalues. Simultaneously, the indices for* L_j *blocks must be changed as* $m_\alpha := m_\alpha + 1$ *for* $\alpha = 0, \ldots, p$, $s_\alpha := s_\alpha + 1$ *for* $\alpha = 0, \ldots, p-1$ *and the indices for* L_j^T *blocks must be changed as* $m_\beta := m_\beta + 1$ *for* $\beta = 0, \ldots, q$, $s_\beta := s_\beta + 1$ *for* $\beta = 0, \ldots, q-1$, *where* $p + q + 1 = \sum_{i=1}^{t} k_i$. *These operations replace the regular part corresponding to the* t *Jordan blocks with a generic square singular pencil* $L_p \oplus L_q^T$.

Similarly, given $\mathcal{P}_1$ it is possible to characterize a pencil $\mathcal{P}_2$ such that $\overline{\text{orbit}}(\mathcal{P}_2) \supset \overline{\text{orbit}}(\mathcal{P}_1)$ and $\mathcal{P}_1$ and $\mathcal{P}_2$ fulfill necessary conditions for two consecutive pencils. Of course, this will impose different prerequisites on $\mathcal{P}_1$'s structure indices. The details are omitted here.

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 [3, 4]. Assuming a fixed accuracy of the input data it is possible to increase or decrease the tolerance for rank decisions such that a more non-generic pencil or a more generic pencil, respectively, is computed. Alternatively, given a Kronecker structure computed by the staircase algorithm we can impose a more non-generic Kronecker structure by applying any of the applicable structure index changes in Corollary 5.1. A modified GUPTRI algorithm can deliver an upper bound on the size of the distance from the pencil $\mathcal{P}_1$ we started with to the pencil $\mathcal{P}_2$ we imposed such that $\mathcal{P}_2$ is in $\overline{\text{orbit}}(\mathcal{P}_1)$. The other way around, we can start with a pencil $\mathcal{P}_1$ and construct a more generic pencil $\mathcal{P}_2$ by adding perturbations (whose sizes depend on the rank decision tolerance) such that $\mathcal{P}_1$ is in $\overline{\text{orbit}}(\mathcal{P}_2)$ and the necessary conditions for consecutive pencils are fulfilled.

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 [7] for computable normwise bounds of the smallest perturbations for going downwards (or upwards) in the closure hierarchy of the set of 2×3 pencils.

6 An Algorithm for Generating Closure Hierarchies

In the following we give an algorithm for deriving the complete closure hierarchy below a given m-by-n pencil $\mathcal{P}$. By choosing $\mathcal{P}$ as the generic pencil, the algorithm derives the complete closure hierarchy for m-by-n pencils. We call $\mathcal{P}_i$ a successor of $\mathcal{P}$ if and only if

$\mathcal{P}$ and $\mathcal{P}_i$ are consecutive and $\overline{\text{orbit}}(\mathcal{P}) \supset \overline{\text{orbit}}(\mathcal{P}_i)$. Algorithm 6.1 generates the closure hierarchy by finding all successors of each KCF.

Algorithm 6.1 *Let $\mathcal{P}$ denote the starting KCF and let Q denote a queue of generated structures, whose successors are not yet determined. Initially Q is empty.*

1. *Generate all structures, $\mathcal{P}_1, \mathcal{P}_2, \ldots$, that are candidates for being successors of $\mathcal{P}$, by applying all appropriate items (i) - (vi) of Corollary 5.1, on all appropriate combination of blocks in $\mathcal{P}$. That is, change the m_i and s_j according to each item where the prerequisites on the indices are fulfilled.*
2. *Compare all pairs $(\mathcal{P}_i, \mathcal{P}_j)$ of candidate successors of $\mathcal{P}$, using Theorem 4.3 (or Theorem 3.1) to detect if $\overline{\text{orbit}}(\mathcal{P}_i) \supset \overline{\text{orbit}}(\mathcal{P}_j)$. If so, the $\mathcal{P}_j$ and $\mathcal{P}$ are not consecutive, and $\mathcal{P}_j$ is discarded from being a successor of $\mathcal{P}$. All other $\mathcal{P}_i$s are registered as successors of $\mathcal{P}$.*
3. *All remaining $\mathcal{P}_i$s that are not already in Q are inserted in order of increasing codimension.*
4. *If Q is not empty, remove the first KCF in Q and denote it $\mathcal{P}$. Repeat from 1.*

We illustrate the execution of Algorithm 6.1 by showing how it generates the complete closure hierarchy for the set of 2-by-3 matrix pencils in Table 1. The algorithm starts with $\mathcal{P} = L_2$, i.e., the generic 2-by-3 Kronecker structure. For each $\mathcal{P}$, the table shows all candidate successors as generated from step 1 in the algorithm. Candidate successors that are discarded in step 2 are overcrossed in Table 1. The graph displaying the complete closure hierarchy in Figure 1 is obtained by drawing arrows from each structure $\mathcal{P}$ in the table, to all its successors (but not to the overcrossed ones).

Table 1: Successors in the closure hierarchy for the set of 2-by-3 matrix pencils, in order as generated by Algorithm 6.1.

$\mathcal{P}$	Successors
L_2:	$L_1 \oplus J_1(\gamma)$
$L_1 \oplus J_1(\gamma)$:	$L_0 \oplus J_1(\gamma) \oplus J_1(\delta)$, $\cancel{L_0 \oplus 2J_1(\gamma)}$, $L_0 \oplus J_2(\gamma)$, $\cancel{L_0 \oplus L_1 \oplus L_1^T}$
$L_0 \oplus J_1(\gamma) \oplus J_1(\delta)$:	$\cancel{2L_0 \oplus J_1(\gamma) \oplus L_0^T}$, $2L_0 \oplus L_1^T$, $L_0 \oplus L_1 \oplus L_1^T$
$L_0 \oplus J_2(\gamma)$:	$L_0 \oplus 2J_1(\gamma)$, $2L_0 \oplus L_1^T$, $L_0 \oplus L_1 \oplus L_1^T$
$L_0 \oplus L_1 \oplus L_1^T$:	$2L_0 \oplus J_1(\gamma) \oplus L_0^T$
$2L_0 \oplus L_1^T$:	$2L_0 \oplus J_1(\gamma) \oplus L_0^T$
$L_0 \oplus 2J_1(\gamma)$:	$2L_0 \oplus J_1(\gamma) \oplus L_0^T$
$2L_0 \oplus J_1(\gamma) \oplus L_0^T$:	$3L_0 \oplus L_0^T$
$3L_0 \oplus L_0^T$:	

To obtain an algorithm for computing the structure hierarchy above a given pencil we have to do the following. In step 1 in Algorithm 6.1, the use of Corollary 5.1 should be replaced by use of a corresponding corollary, that given a pencil $\mathcal{P}_1$ gives the prerequisites and the required changes in the m_i and s_j indices for finding a pencil $\mathcal{P}_2$ such that $\overline{\text{orbit}}(\mathcal{P}_2) \supset \overline{\text{orbit}}(\mathcal{P}_1)$ and $\mathcal{P}_1$ and $\mathcal{P}_2$ fulfill necessary conditions for consecutive pencils, according to Lemma 3.1. In step 2, $\mathcal{P}_i$ should be discarded from being successor of $\mathcal{P}$ if $\overline{\text{orbit}}(\mathcal{P}_i) \supset \overline{\text{orbit}}(\mathcal{P}_j)$. In step 3, Q should be sorted in decreasing order of codimension. Finally, to derive the complete closure hierarchy, we should now start with the most non-generic pencil, i.e., the zero pencil.

We now from [7] that that number of different Kronecker structures grows rapidly for increasing m and n. In Algorithm 6.1, this will not only cause a lot of more work for generation of all true successors, but there will also be a large growth of the number of generated candidate successors that in step 2 are discarded. Therefore, the amount of redundant work in the algorithm will increase with increasing m and n.

7 Conclusions

We have given new interpretations of important results by Pokrzywa [11], for determining closure relations among orbits of Kronecker structures. This has partly been done by generalizing classical theorems by Gantmacher [8]. These results have been used to derive an algorithm for computation of the complete Kronecker structure hierarchy, or the complete hierarchy above or below a given Kronecker structure. The algorithm is presented in terms of the rank-decisions required in a staircase algorithm, in order to compute the Kronecker structure hierarchy.

From a mathematical point of view, we know that the Kronecker structures that are closest to a pencil $\mathcal{P}$ with a given structure are the ones that are above it in the hierarchy. We know that these structures are the ones that are found by adding perturbations in the normal space of orbit($\mathcal{P}$) in the point given by $\mathcal{P}$. How to actually compute these perturbations, or the *versal deformation* of a KCF, is shown in [6]. For our purposes, it is not the versal deformation itself that is most important. We are more interested in metrical information for the perturbation theory relevant to the Kronecker canonical form. The results in this contribution should be seen in this context and as preliminary results for the second part of [6].

Acknowledgements

I would like to thank Bo Kågström for helpful discussions and constructive comments during the work and the preparation of this paper. I would also like to thank Alan Edelman for helpful discussions and for encouraging the study of the stratification of the Kronecker canonical form.

References

[1] T. Beelen and P. Van Dooren. An improved algorithm for the computation of Kronecker's canonical form of a singular pencil. *Lin. Alg. Appl.*, 105:9–65, 1988.

[2] J. Demmel and A. Edelman. The Dimension of Matrices (Matrix Pencils) with Given Jordan (Kronecker) Canonical Forms. Report LBL-31839, Mathematics Department, Lawrence Berkeley Laboratories, University of California, Berkeley, CA 94720, 1992. To appear in *Lin. Alg. Appl.*

[3] 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*, Vol.19(No. 2):160–174, June 1993.

[4] 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*, Vol.19(No. 2):175–201, June 1993.

[5] C.C. Mac Duffee. *The Theory of Matrices.* Chelsea Publishing Company, New York, 1956.

[6] A. Edelman, E. Elmroth, and B. Kågström. A Geometric Approach To Perturbation Theory of Matrices and Matrix Pencils. Part I: Versal deformations. Report UMINF-95.09, Department of Computing Science, Umeå University, S-901 87 Umeå, Sweden, March, 1995.

[7] E. Elmroth and B. Kågström. The Set of 2-by-3 Matrix Pencils – Kronecker Structures and Their Transitions Under Perturbations. Report UMINF-93.22, Department of Computing Science, Umeå University, S-901 87 Umeå, Sweden, November, 1993. Revised October, 1994. To appear in *SIAM J. Matrix Anal. Appl.*

[8] F. Gantmacher. *The Theory of Matrices, Vol. I and II (transl.).* Chelsea, New York, 1959.

[9] B. Kågström. RGSVD - an algorithm for computing the Kronecker canonical form and reducing subspaces of singular matrix pencils $A - \lambda B$. *SIAM J. Sci. Stat. Comp.*, 7(1):185–211, 1986.

[10] V. B. Khazanov and V. Kublanovskaya. Spectral problems for matrix pencils. Methods and algorithms. I. *Sov. J. Numer. Anal. Math. Modelling*, 3:337–371, 1988.

[11] A. Pokrzywa. On perturbations and the equivalence orbit of a matrix pencil. *Lin. Alg. Appl.*, 82:99–121, 1986.

[12] P. Van Dooren. The computation of Kronecker's canonical form of a singular pencil. *Lin. Alg. Appl.*, 27:103–141, 1979.

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

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