A Geometric Approach to Perturbation Theory of Matrices and Matrix Pencils. Part I: Versal Deformations

Alan Edelman*, Erik Elmroth†, and Bo Kågström†

March 23, 1995
UMINF-95.09
ISSN-0348-0542

Abstract

We derive versal deformations of the Kronecker canonical form by deriving the tangent space and orthogonal bases for the normal space to the orbits of strictly equivalent matrix pencils. These deformations reveal the local perturbation theory of matrix pencils related to the Kronecker canonical form. We also obtain a new singular value bound for the distance to the orbits of less generic pencils. The concepts, results and their derivations are mainly expressed in the language of numerical linear algebra. We conclude with experiments and applications.

Keywords: Jordan canonical form, Kronecker canonical form, generalized Schur decomposition, staircase algorithm, versal deformations, tangent and normal spaces, singularity theory, perturbation theory.

---

*Department of Mathematics, Room 2–380, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, Email address: edelman@math.mit.edu, Supported by NSF grant DMS-9120852 and an Alfred P. Sloan Foundation Research Fellowship.

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

Contents

1 Introduction and Examples

1.1 Introduction

Traditionally, canonical structure computations take as their input some mathematical object, a matrix or a pencil say, and return an equivalent object that is perhaps simpler, or makes clear the structure of the equivalence relation. Some example equivalence relations and corresponding canonical forms are:

Structure	**Equivalence Relation**	**Canonical Form**
Square Matrices	$A \sim X^{-1}AX$	Jordan Canonical Form
Rectangular Matrices	$A \sim UAV$	Singular Values
Rectangular Matrices	$A \sim XA$	Reduced Echelon form
Matrix Pencils	$A - \lambda B \sim P^{-1}(A - \lambda B)Q$	Kronecker Canonical Form
Analytic real functions	$f(x) \sim f(\phi(x))$	$\pm x^k$

In the first three examples the input is a matrix, in the next example, the input is a pencil. In these cases, X, P, and Q are presumed nonsingular, and U and V are presumed orthogonal. We presume the real functions f are analytic in a neighborhood of zero, $f(0) = 0$, $\phi(0) = 0$ and $\phi(x)$ is monotonic and analytic near zero.

Canonical forms appear in every branch of mathematics. A few examples from control theory may be found in [20, 19, 25, 18]. However, researchers in singularity theory have asked the question what happens if you have not one object that you want to put into a normal form, but rather a whole family of objects nearby some particular object and you wish to put each member of the family into a canonical form in such a way that the canonical form depends smoothly on the deformation parameters.

For example, one may have, a one parameter matrix deformation of A_0 which is simply an analytic function $A(\lambda)$ for which $A(0) = A_0$. An n parameter deformation is defined the same way, except that $\lambda \in \mathbf{R}^n$. Similarly, one may have n parameter deformations of pencils or functions. Sticking with the matrix example, we say two deformations $A(\lambda)$ and $B(\lambda)$ are equivalent if $A(\lambda)$ and $B(\lambda)$ have the same Jordan canonical form for each and every λ. A deformation of a matrix is said to be versal if, loosely speaking, it captures all possible Jordan form behaviors, near the matrix. A deformation is said to be miniversal, if it does so with as few parameters as possible. A more formal discussion of these definitions may be found in Section 2.

Derivation of versal and miniversal deformations requires a detailed understanding of the perturbation theory of the objects under study. In particular, one needs to understand the tangent space of the equivalence relation, and how it is embedded in the entire space. In Section 2, we explain the mechanics of this perturbation theory.

While we believe that versal deformations are interesting mathematical objects, this work differs from other works on the subject in that our primary goal is not so much the versal deformation or the miniversal deformation, but rather the perturbation theory and how it influences the computation of the Kronecker canonical form. As such we tend to be interested more in metrical information than topological information. Therefore, we obtain new distance formulas to the space of less generic matrix pencils in Section 4. In Section 5, we derive an explicit orthogonal basis for the normal space of a Kronecker canonical form. For us a versal decomposition will be an explicit decomposition of a perturbation into its tangential and normal components, and we will not derive any miniversal deformations that may have simpler forms, but hide the metric information.

Versal deformations for function spaces are discussed in [17, 24, 4, 5]. The first application of these ideas for the matrix Jordan canonical form is due to Arnold [1]. Further references closely related to Arnold's matrix approach are [28] and [6]. The latter reference, [6], also includes applications to differential equations. Applications of the matrix idea towards an understanding of companion matrix eigenvalue calculations may be found in [13]. The only other work that we are aware of that considers versal deformations of the Kronecker canonical form is by Berg and Kwatny [3] who have independently derived some of the normal forms considered in this paper.

Our Section 2 contains a thorough explanation of versal deformations from a linear algebra perspective. Chapter 3 briefly reviews matrix pencils and canonical forms. Chapter 4 derives the geometry of the tangent and normal spaces to the orbits of matrix pencils. Chapter 5 derives the versal deformations, while Chapter 6 gives applications and illustrations.

1.2 Geometry of matrix space

Our guiding message is very simple: matrices should be seen in the mind's eye geometrically as points in n^2 dimensional space. A perfect vision of numerical computation would allow us to picture computations as moving matrices from point to point or manifold to manifold.

Abstractly, it hardly matters whether a vector is a column of numbers or a geometric point in space. However, without the interplay of these two representations, numerical linear algebra would not be the same. Imagine explaining how Householder reflections transform vectors without the geometric viewpoint.

By contrast, in numerical linear algebra we all know that matrices are geometric points in n^2 dimensional space, but it is far rarer that we actually *think* about them this way. Most often, matrices are thought of as either (sparse or dense) arrays of numbers, or they are operators on vectors.

The Eckart–Young (or Schmidt–Mirsky theorem) [27, p.210] gives a feel for the geometric approach. The theorem states that the smallest singular value of A is the Frobenius distance of A to the set of singular matrices. One can not help but to see a blob representing the set of singular matrices. This amorphous blob is most often thought of as an undesirable part of town, so unfortunately numerical analysts hardly ever study the set itself.

Demmel has helped to pioneer the development of geometric techniques [7] for the analysis of ill-conditioning of numerical analysis problems. Shub and Smale [26] are applying geometrical approaches towards the solution of polynomial systems.

We believe that if only we could better understand the geometry of matrix space, our knowledge of numerical algorithms and their failures would also improve. A general program for numerical linear algebra, then, is to transfer from pure mathematicians the technology to understand geometrically the high dimensional objects that arise in numerical linear algebra. This program may not be easy to follow. A major difficulty is that pure mathematicians pay a price for their beautiful abstractions – they do not always possess a deep understanding of the individual objects that we wish to study. This makes technology transfer difficult. Even when the understanding exists somewhere, it may be difficult to recognize or may be buried under a heavy layer of notation. This makes technology transfer time consuming. Finally, even after putting in the time for the excavation, the knowledge may still be difficult to apply towards the understanding or the improving of practical algorithms. This makes technology transfer from pure mathematics frustrating.

Nevertheless, our goal as researchers is the quest for understanding which we may then apply. In this paper, we follow our program for the understanding of the Jordan and Kronecker canonical forms of matrices and matrix pencils, respectively. Many of the ideas to be found in this paper have been borrowed from the pure mathematics literature with the goal of simplifying and applying to the needs of numerical linear algebraists.

While this is quite a general program for numerical linear algebra, this paper focuses on a particular goal. We analyze *versal deformations* from the numerical linear algebra viewpoint, and then compute normal deformations for the Kronecker canonical form. We consider both of these as stepping stones towards the far more difficult goal of truly understanding and improving upon staircase algorithms for the Jordan or Kronecker canonical form. These are algorithms used in systems and control theory. The structures of these matrices or pencils reflect important physical properties of the systems they model, such as controllability [10, 30].

The user chooses a parameter η to measure any uncertainty in the data. The existence of a matrix or pencil with a different structure within distance η of the input means that the actual system may have a different structure than the approximation supplied as input. These algorithms try to perturb their input by at most η so as to find a matrix or pencil with as high a codimension as possible. The algorithm is said to *fail* if there is another perturbation of size at most η which would raise the codimension even further. Therefore, we need to understand the geometry of matrix space in order to begin to understand how we can supply the correct information to the user. With this information, we believe that we would then be able to not only correctly provide the least generic solutions, but also understand how singularities hinder this process. Bad solutions may then be refined so as to obtain better solutions. As the next subsection illustrates, the geometry directly affects the perturbation theory.

1.3 Motivation: a singular value puzzle

Consider the following four nearly singular matrices:

$$M_1 = \begin{pmatrix} 0 & 1+\epsilon \\ 0 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 1 \\ \epsilon & 0 \end{pmatrix}, \quad M_3 = \begin{pmatrix} \epsilon & 1 \\ 0 & -\epsilon \end{pmatrix}, \quad M_4 = \begin{pmatrix} \epsilon & 1 \\ 0 & \epsilon \end{pmatrix}. \qquad (1.1)$$

Each of these matrices are distance $O(\epsilon)$ from the Jordan block $J_2(0) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. What is the smaller of the two singular values of each of M_1, M_2, M_3 and M_4? The answer is

$$\sigma_{\min}(M_1) = 0, \quad \sigma_{\min}(M_2) = \epsilon, \quad \sigma_{\min}(M_3) \approx \epsilon^2 \quad \text{and} \quad \sigma_{\min}(M_4) \approx \epsilon^2.$$

A quick way to verify this algebraically is to notice that the larger singular value of each matrix is approximately 1 so that the smaller is approximately the (absolute) determinant of the matrix. Another approach that bounds the smallest singular value is the combination of the Eckart–Young theorem and the observation that these matrices are singular:

$$M_1' = M_1, \qquad M_2' = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \; M_3' = \begin{pmatrix} \epsilon & 1 \\ -\epsilon^2 & -\epsilon \end{pmatrix}, \; M_4' = \begin{pmatrix} \epsilon & 1 \\ \epsilon^2 & \epsilon \end{pmatrix}.$$

When $\epsilon = 0$ in (1.1) our four matrices become the singular 2×2 Jordan block $J_2(0)$. As ϵ varies from 0 each of the four forms in (1.1) traces out a line in matrix space. The geometric issue that is interesting here is that the line of matrices traced out as ϵ varies is { 1:In 2:Normal 3:Tangent 4:Tangent } to the set of singular matrices. Somehow, this feels like the "right" explanation for why the smaller singular values are { 1:0, 2:ϵ, 3:$\approx \epsilon^2$, 4:$\approx \epsilon^2$ }.

Let us take a closer look at the set of singular matrices. The four parameters found in a 2×2 matrix M are best viewed in a transformed coordinate system:

$$\begin{aligned} M = (x, y, z, w) &= x \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + w \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} w + z & x \\ y & w - z \end{pmatrix}. \end{aligned}$$

In this coordinate system, the singular matrices fall on the surface described by the equation $w^2 = z^2 + xy$. This is a three dimensional surface in four dimensional space. The traceless singular matrices ($w = 0$) fall on the cone $z^2 + xy = 0$ in three dimensional space.

Our matrix $J_2(0)$ may now be represented as $(1, 0, 0, 0)$ and the four lines of matrices mentioned above are

$$\begin{aligned} l_1 = \{ \quad (1 + x, 0, 0, 0) \quad \} &= \left\{ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \epsilon \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right\}, \\ l_2 = \{ \quad (1, y, 0, 0) \quad \} &= \left\{ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \epsilon \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\}, \\ l_3 = \{ \quad (1, 0, z, 0) \quad \} &= \left\{ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \epsilon \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}, \\ l_4 = \{ \quad (1, 0, 0, w) \quad \} &= \left\{ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \epsilon \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}. \end{aligned}$$

The lines l_1, l_2 and l_3 are all traceless, i.e., the matrices on each of these lines may be viewed in the three dimensional space of the cone. The line l_1 is not only tangent to the cone, but in fact it lies in the cone. The line l_3 is tangent to one of the circular cross-sections of the cone.

Figure 1 illustrates l_3 as a "stick" resting near the bottom of the cone. The line l_1 is a thin line on the cone through the same point.

The line l_4 is normal to the cone but it is also tangent to the manifold of singular matrices. One way to picture this in three dimensions is to take the three dimensional slice

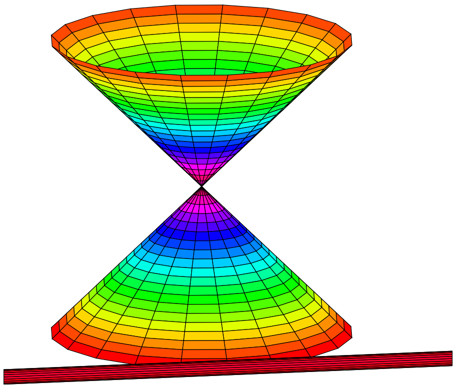

Figure 1: Cone of traceless singular matrices with “stick” representing a tangent

of $\{w^2 = z^2 + xy\}$ corresponding to $x = 1$, i.e., $\{w^2 - z^2 = y\}$. This is a hyperboloid with the Jordan block as a saddle point. The line is the tangent to the parabola $w^2 = y$ which rests in the plane $z = 0$. Figure 2 illustrates this line with a cylindrical stick whose central axis is the tangent. Lastly, the line l_2 is normal to the set of singular matrices.

If we move a distance ϵ away from a point on a surface along a tangent, our distance to the surface remains $O(\epsilon^2)$. This is what the singular value corresponding to l_3 and l_4 is telling us. Alternatively, if we move normal to the surface as in l_2, the singular value changes more rapidly: $O(\epsilon)$.

The cone of singular matrices with $w = 0$ is not only a slice of a large dimensional space, but it is also the (closure of) the set of matrices similar to $J_2(0)$ (which we denote $\text{orbit}(J_2(0))$ in Section 2.4). The matrices similar to $J_2(0)$ are singular and traceless. In fact, the only matrix that is singular and traceless that is not similar to $J_2(0)$ is the 0 matrix which is the vertex of the cone. We further explore this case in Section 2.5 after we have defined versal deformations.

We conclude that the geometry of the orbit and in particular the directions of the tangents and normals to the orbit directly influence the eigenvalue perturbation theory.

2 Introduction to Versal Deformations

This introduction is designed to be readable for general audiences, but we particularly target the numerical linear algebra community.

The ideas here may be thought of as a numerical analyst's viewpoint on ideas that were inspired by Arnold's work [1] on versal deformations of matrices. Further elaboration upon Arnold's versal deformations of matrices may be found in [6, Chapter 2.9 and 2.10] and [28]. These ideas fit into a larger context of differential topology and singularity theory. Bruce and Giblin [5] have written a wonderfully readable introduction to singularity theory emphasizing the elementary geometrical viewpoint. After reading this introduction, it is easy to be lulled into the belief that one has mastered the subject, but a whole further more advanced wealth of information may be found in [17, 24, 4]. Finally, what none of these

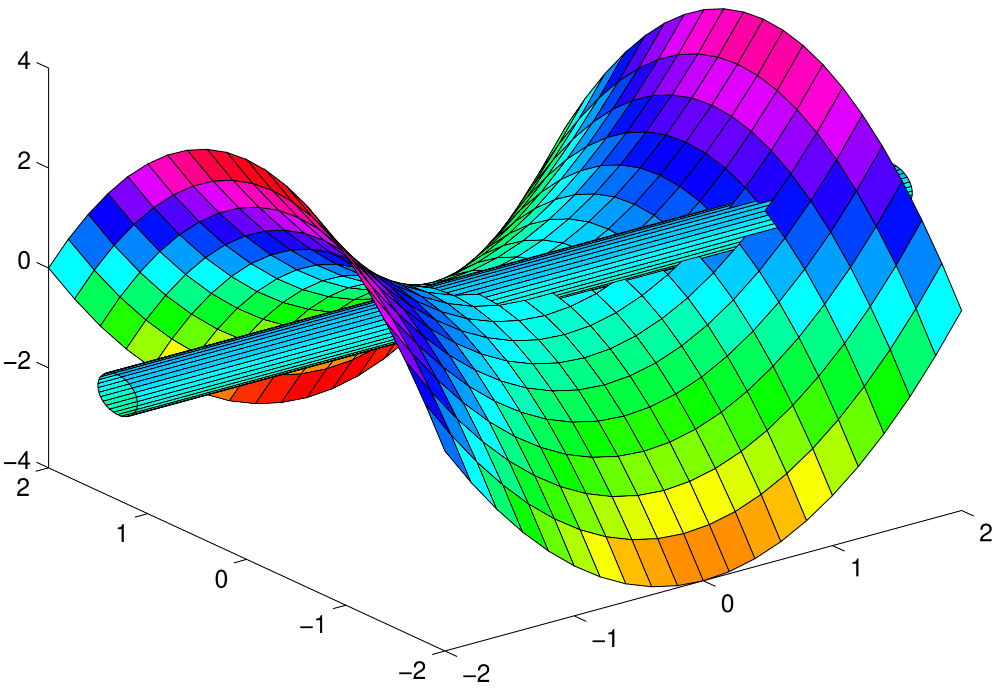

Figure 2: Manifold of singular matrices. The axis of the cylindrical stick is tangent to the manifold.

references do very well is explain clearly that there is still much in this area that mankind does not yet fully understand.

Singularity theory may be viewed as a branch of the study of curves and surfaces, but its crowning application is towards the topological understanding of functions and their behavior under perturbations. Of course, numerical analysts are very interested in perturbations as well.

2.1 Characteristic polynomials give the "feel" of versal deformations

Let $A(\lambda)$ be a differentiable one parameter family of matrices through $A_0 \equiv A(0)$. This is just a curve in matrix space. If A_0 has a complicated Jordan canonical form, then very likely, the Jordan canonical form of $A(\lambda)$ is a discontinuous function of λ. (The Jordan canonical form, you will remember, can have nasty ones popping up unexpectedly on the superdiagonal.) It is even more desirable if that function can somehow describe the kinds of matrices that are near A_0.

Discontinuities are as unpleasant for pure mathematicians as they are for computers. Therefore Arnold [1] asks what kinds functions of λ are differentiable? (or many times differentiable, or analytic.)

One function that comes to mind is the characteristic polynomial $p_\lambda(t) \equiv \det(A(\lambda) - tI)$. The coefficients of p_λ are clearly differentiable functions of λ no matter how complicated a Jordan canonical form the matrix A_0 might have. In numerical linear algebra, we never compute the characteristic polynomial, because the eigenvalues are often very poorly determined by the coefficients of the characteristic polynomial. Mathematically, the characteristic polynomial is a nice function of a matrix because its coefficients, unlike the eigenvalues of the matrix, are analytic functions of the entries of the matrix.

The characteristic polynomial is a reasonable representation for the Jordan canonical form under the special circumstance that every matrix $A(\lambda)$ is *non-derogatory* (i.e., each matrix has exactly one Jordan block for each distinct eigenvalue). By a reasonable representation, we mean here that it actually encodes the Jordan canonical form of A_0. Theoretically, if you know the characteristic polynomial, then you know the eigenvalues with appropriate multiplicities. It follows that there is a unique non-derogatory Jordan canonical form. (See Wilkinson [33, pp.11–16 or Note 55, p.408]). To repeat, there is a one-to-one correspondence among the n eigenvalues of a non-derogatory matrix, the characteristic polynomial of a non-derogatory matrix, and the Jordan canonical form of a non-derogatory matrix, but only the characteristic polynomial is a differentiable function of the perturbation parameter λ. (The eigenvalues themselves can have first order perturbations with the non-differentiable form $\lambda^{1/n}$, for example, for an $n \times n$ matrix A_0 with only one Jordan block $J_n(\lambda)$. This is a well known example.)

In the language of numerical linear algebra, we would say that a non-derogatory matrix A_0 may be written in companion matrix form KCK^{-1}, in such a way that differentiable perturbations to the matrix A_0 lead to differentiable perturbations to the companion matrix C. Here the matrix K is a Krylov matrix. (See [16, p. 369]). Equivalently, first order perturbations to the matrix A_0 are manifested as first order perturbations to the companion matrix C. When A_0 is a companion matrix, this gives a first order perturbation theory for the characteristic polynomials of nearby matrices. This perturbation theory is computed in

[13].

Our story would almost stop here if we were only interested in the Jordan form of non-derogatory matrices. We use "almost" because it would be a shame to stop here without explaining the ideas geometrically. Even if we did not discuss the geometry, we have reasons to continue on, since matrix space is enriched with the derogatory matrices, and also we wish to generalize these ideas about the Jordan canonical form to cover the more complicated case of the Kronecker canonical form.

2.2 The rational canonical form is not enough for derogatory matrices

In the previous subsection we saw that n parameters were sufficient to specify the Jordan canonical form of any matrix in a small neighborhood of a non-derogatory matrix. What happens if the matrix is derogatory? One obvious guess turns out to be wrong. The usual generalization of the companion matrix form for derogatory matrices is the rational canonical form. If A_0 is derogatory, it may be put in rational canonical form. This form may be thought of as the direct sum of companion matrices C_i with dimension $m_1 \geq m_2 \geq \ldots \geq m_k$. The characteristic polynomial of each C_i divides the characteristic polynomial of all the preceding $C_j, j < i$. Can any nearby matrix be expressed as the direct sum of companion matrices with dimension $m_1, m_2, \ldots, m_k$ in a nice differentiable manner? The answer is generally no; though good enough to specify the Jordan canonical form of a matrix, the rational canonical form fails to be powerful enough to specify the Jordan canonical forms of all matrices in a neighborhood. The reason is that there are just not enough parameters in the rational canonical form to cover all the possibilities. To have enough parameters we need a "versal deformation".

2.3 Versal deformation: the linearized theory

The "linearized" picture of a versal deformation is easy to understand. We therefore explain this picture before plunging into the global point of view. The general case may be nonlinear, but the linearized theory is all that really matters. For simplicity we assume that we are in real n dimensional Euclidean space, but this assumption is not so important.

We recall the elementary fact that if $\mathcal{S}$ and $\mathcal{T}$ are subspaces of $\mathbf{R}^n$ such that $\mathcal{S}+\mathcal{T} = \mathbf{R}^n$, then there exist linear projections $\pi_{\mathcal{S}}$ and $\pi_{\mathcal{T}}$ that map onto $\mathcal{S}$ and $\mathcal{T}$, respectively.

Consider a point $x \in \mathcal{S}$. We will investigate all possible perturbations y of x, but we will not be concerned with perturbations that are within $\mathcal{S}$ itself. Psychologically, we consider all the vectors in $\mathcal{S}$ to somehow be the same so there will be no need to distinguish them. Let $\mathcal{T}$ be any linear subspace such that $\mathcal{S} + \mathcal{T} = \mathbf{R}^n$, i.e., any vector may be written as the sum of an element of $\mathcal{T}$ and an element of $\mathcal{S}$ (not necessarily uniquely). Clearly if $t_1, \ldots, t_k$ span $\mathcal{T}$, then our perturbed vector $x + y$ may be written as

$$x + y = x + \sum_{i=1}^{k} \lambda_i t_i + (\text{something in } \mathcal{S}),$$

where the λ_i may be chosen as linear functions of y. We see here what will turn out to be the key idea of a versal deformation, every perturbation vector may be expressed in terms of the λ_i and vectors that we are considering to all be equivalent.

We now formally introduce the local picture of versal deformations.

Definition 2.1 *A* linear deformation *of the point x is a function defined on $\lambda \in \mathbf{R}^l$:*

$$A(\lambda) = x + T\lambda,$$

where $T = [t_1 t_2 \ldots t_l]$ are arbitrary directions.

The choice of the word "deformation" is meant to convey the idea that we are looking at small values of the λ_i, and these perturbations are small deformations of the starting point x.

Definition 2.2 *A linear deformation $A(\lambda)$ of the point x is* versal *if for all linear deformations $B(\mu)$ of the point x, it is possible to write*

$$B(\mu) = A(\phi(\mu)) + \theta(\mu),$$

where $\phi(\mu)$ is a linear function from $\mu_1, \ldots, \mu_m$ to $\lambda_1, \ldots, \lambda_l$ with $\phi(0) = 0$, and θ is a linear function from μ into $\mathcal{S}$, with $S(0) = 0$.

We now explain why $A(\lambda) = x + \sum_{i=1}^{l} \lambda_i t_i$ is versal if and only if $\mathcal{S} + \mathcal{T} = \mathbf{R}^n$. Clearly $A(\phi(\mu)) + \theta(\mu) \in \mathcal{S} + \mathcal{T}$ and since B may be arbitrary, it is necessary that $\mathrm{span}(\{t_i\}) + \mathcal{S} = \mathbf{R}^n$. It is also sufficient, because we then obtain linear projections allowing us to write $B(\mu) = x + \pi_{\mathcal{S}} B(\mu) + \pi_{\mathcal{T}} B(\mu)$. The functions ϕ and θ may be obtained from $\pi_{\mathcal{S}}$ and $\pi_{\mathcal{T}}$.

Definition 2.3 *A linear deformation $A(\lambda)$ of the point x is* universal *or* miniversal *if it is versal, and has the fewest possible parameters needed for a versal deformation.*

The number of parameters in a miniversal deformation is exactly the codimension of $\mathcal{S}$. Numerical analysts might prefer taking the t_i to be an orthogonal basis for $\mathcal{S}^{\perp}$, the subspace perpendicular to $\mathcal{S}$. This provides one natural miniversal deformation. Arnold [1] does not insist on using $\mathcal{S}^{\perp}$, any basis for any subspace of dimension $n - \dim \mathcal{S}$ will do provided that it intersects $\mathcal{S}$ at zero only. From the topological point of view, this is exactly the same, though of course the numerical properties may be quite different.

2.4 Versal deformations the bigger picture

The previous subsection explained the linear or first order theory of versal deformations. At this point, the reader might wonder whether this is just a whole lot of jargon to merely extend a basis for a subspace to the entire space. At the risk of delaying the motivation until now, we decided to make sure that the linear theory is well understood.

We are still in a finite dimensional Euclidean space $\mathbf{R}^n$, but $\mathcal{S}$ will no longer be a flat subspace. Instead, we wish to consider any equivalence relation $\sim$, such that the *orbit* of x ($\mathrm{orbit}(x) \equiv \{y | y \sim x\}$) is a smooth submanifold. As an example we might define $x \sim y$ to mean $\|x\| = \|y\|$, in which case the orbits are spheres. In this context the word "orbit" is quite natural. In n^2 dimensional space, points may be thought of as $n \times n$ matrices, and the orbit is the set of matrices with the same Jordan canonical form.

One final example that we must mention (because it explains the origins and significance of singularity theory) lives in an infinite dimensional space. The vector space is the set of analytic functions $f(x)$ for which $f(0) = 0$. We can define $f \sim g$, if $f(x)$ and $g(\phi(x))$ have the same Taylor expansion at $x = 0$, where ϕ is a monotonic analytic function with $\phi(0) = 0$. The orbit of any function is some complicated infinite dimensional manifold, but the codimension of the manifold happens to be finite.

Returning to $\mathbf{R}^n$, we can now cast everything into a nonlinear context.

Definition 2.4 *A* deformation *of the point x is any differentiable function*

$$A(\lambda_1, \ldots, \lambda_l)$$

satisfying $A(0) = x$.

Definition 2.5 *A deformation $A(\lambda)$ of the point x is* versal *if for all deformations $B(\mu)$, it is possible to write*

$$B(\mu) \sim A(\phi(\mu))$$

in an arbitrarily small neighborhood of 0, where $\phi(\mu)$ is a differentiable function from $\mu_1, \ldots, \mu_m$ to $\lambda_1, \ldots, \lambda_l$ for which $\phi(0) = 0$.

The good news is that the inverse function theorem lets us express this nonlinear notion in terms of the linear theory:

Theorem 2.1 *A deformation $A(\lambda)$ of x is versal if and only if $A_*(\lambda)$ is a linear deformation at the point x on the subspace* $\tan(\text{orbit}(x))$*, where A_* is the linearization of A near x (i.e. only first derivatives matter), and* tan *denotes the subspace tangent to the orbit at x.*

The rigorous proof may be found in [1], but the intuition should be clear: near the point x, only linear deformations matter, and the curvature of the orbit becomes unimportant: only the tangent plane matters. In other words $y \sim x$ only if y is in the orbit of x, but to first order, $y \sim x$ if (roughly speaking) $y = x + s$, where s is a small tangent vector to the orbit.

2.5 Versal deformations for the Jordan canonical form

We begin with deformations of the matrix $A_0 = J_2(0)$. The perturbation theory and the normal and tangent spaces were discussed in Section 1.3. We will use the same coordinate system here.

Four parameters $\mu = (\mu_1, \mu_2, \mu_3, \mu_4)$ are sufficient to describe the most general deformation of A_0:

$$A(\mu) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \mu_1 & \mu_2 \\ \mu_3 & \mu_4 \end{pmatrix}.$$

The equivalence relation is that of similar matrices, and it is easy to see by checking the trace and determinant, that for sufficiently small values of μ, we have the equivalence,

$$A(\mu) \sim B(\lambda) \equiv \begin{pmatrix} 0 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix},$$

where $\lambda = \phi(\mu)$ is defined by $\lambda_1 = \mu_3(1+\mu_2) - \mu_1\mu_4$ and $\lambda_2 = \mu_1 + \mu_4$. It is worth emphasizing that the equivalence relation does not work if $A(\mu)$ is derogatory, but this does not happen for small parameters μ.

We then see from Definition 2.5, that the two parameter deformation $B(\lambda)$ is versal. In fact, it is *miniversal*, in that one needs the two parameters. From the local theory pictured in Section 1.3, we saw that the orbit of $J_2(0)$ is the two dimensional cone, and therefore the tangent and normal spaces are each two dimensional. The number of parameters in a miniversal deformation is always the dimension of the normal space.

It is a worthwhile exercise to derive the similarity transformation $C(\mu)$ for which

$$A(\mu) = C(\mu)^{-1}B(\phi(\mu))C(\mu),$$

and then linearize this map for small values of μ to see which directions fall along the tangent space to the cone, and which directions are normal to the cone.

Now consider deformations of $A_0 = I_2$ or $A_0 = 0$. Both matrices are derogatory with 2 eigenvalues 1 and 0, respectively. The tangent space does not exist (i.e., it is zero dimensional). Any possible behavior may be found near I_2 (or 0) including a one dimensional space of derogatory matrices. The miniversal deformation of I_2 (or 0) is the full deformation requiring four parameters.

The general case has been worked out by Arnold [1]. The tangent vectors to the orbit of a matrix A_0 are those matrices that may be expressed as $XA_0 - A_0X$. The normal space is the adjoint of the centralizer, i.e., the set of matrices Z satisfying

$$A_0^H Z = Z A_0^H.$$

Let A_0 has p distinct eigenvalues $\lambda_i, i = 1 : p$ with p_i Jordan blocks each. Let $q_1(\lambda_i) \geq q_2(\lambda_i) \geq \ldots \geq q_{p_i}(\lambda_i)$ denote the sizes of the Jordan blocks corresponding to the eigenvalue λ_i. Then the dimension of the normal space of A_0 is

$$\sum_{i=1}^{p}\sum_{j=1}^{p_i}(2j-1)q_j(\lambda_i) = \sum_{i=1}^{p}(q_1(\lambda_i) + 3q_2(\lambda_i) + 5q_3(\lambda_i) + \ldots).$$

Notice that the values of the distint λ_i play no role in this formula. The dimension of the normal space of A_0 is determined only by the sizes of the Jordan blocks of A_0 associated with distinct eigenvalues. If the matrix is in Jordan canonical form, then the normal space consists of matrices made up of Toeplitz blocks, whose block structure is completely determined by the sizes of the Jordan blocks for different eigenvalues. The normal space is the same for all matrices with the same Jordan structure independent of the values of the distinct eigenvalues, so one may as well consider only Jordan blocks corresponding to a 0 eigenvalue. This form of the normal space for the zero eigenvalues is a special case in Theorem 5.1.

3 The Algebra of Matrix Pencils – Canonical Forms

We saw in Section 2.4 that to consider versal deformations, one needs a finite or infinite dimensional space, and an equivalence relation on this space. For the remainder of this paper, we consider the finite dimensional Euclidean space of matrix pencils endowed with the Euclidean metric (usually denoted the Frobenius metric in this context). The equivalence relation is that of the strict equivalence of pencils.

We consider a matrix pencil $A - \lambda B$, where A and B are arbitrary $m \times n$ matrices with real or complex entries. The pencil is said to be *regular* if $m = n$ and $\det(A - \lambda B)$ is not identically zero. Indeed, the zeros of $\det(A - \lambda B) = 0$ are the (generalized) eigenvalues of a regular pencil. Otherwise, i.e., if $\det(A - \lambda B)$ is identically zero or $m \neq n$, $A - \lambda B$ is called *singular*. Two $m \times n$ pencils $A_1 - \lambda B_1$ and $A_2 - \lambda B_2$ are *strictly equivalent* if there exist constant (independent of λ) invertible matrices P of size $m \times m$ and Q of size $n \times n$ such that

$$P^{-1}(A_1 - \lambda B_1)Q = A_2 - \lambda B_2.$$

Kronecker has shown that any matrix pencil is strictly equivalent to a canonical diagonal form that describes the structure elements of $A - \lambda B$ (including generalized eigenvalues and eigenspaces) in full detail (e.g. see [15]). This form is a generalization of the Jordan canonical form (JCF) to general matrix pencils.

3.1 Kronecker canonical form

The *Kronecker canonical form* (KCF) of $A - \lambda B$ exhibits the fine structure elements, including elementary divisors (Jordan blocks) and minimal indices, and is defined as follows [15]. Suppose $A, B \in \mathbf{C}^{m\times n}$. Then there exist nonsingular $P \in \mathbf{C}^{m\times m}$ and $Q \in \mathbf{C}^{n\times n}$ such that

$$P^{-1}(A - \lambda B)Q = S - \lambda T, \tag{3.1}$$

where $S = \text{diag}\ (S_{11}, \ldots, S_{bb})$ and $T = \text{diag}\ (T_{11}, \ldots, T_{bb})$ are block diagonal. $S_{ii} - \lambda T_{ii}$ is $m_i \times n_i$. We can partition the columns of P and Q into blocks corresponding to the blocks of $S - \lambda T$: $P = [P_1, \ldots, P_b]$ where P_i is $m \times m_i$, and $Q = [Q_1, \ldots, Q_b]$ where Q_i is $n \times n_i$. Each block $M_i \equiv S_{ii} - \lambda T_{ii}$ must be of one of the following forms: $J_j(\alpha)$, N_j, L_j or L_j^T. First we consider

$$J_j(\alpha) \equiv \begin{bmatrix} \alpha-\lambda & 1 & & \\ & \cdot & \cdot & \\ & & \cdot & 1 \\ & & & \alpha-\lambda \end{bmatrix} \quad \text{and} \quad N_j \equiv \begin{bmatrix} 1 & -\lambda & & \\ & \cdot & \cdot & \\ & & \cdot & -\lambda \\ & & & 1 \end{bmatrix} \tag{3.2}$$

$J_j(\alpha)$ is simply a $j \times j$ Jordan block, and α is called a *finite eigenvalue*. N_j is a $j \times j$ block corresponding to an *infinite eigenvalue* of multiplicity j. The $J_j(\alpha)$ and N_j blocks together constitute the *regular structure* of the pencil. All the $S_{ii} - \lambda T_{ii}$ are regular blocks if and only if $A - \lambda B$ is a regular pencil. $\sigma(A - \lambda B)$ denotes the eigenvalues of the regular part of $A - \lambda B$ (with multiplicities), and is called the *spectrum* of $A - \lambda B$.

The other two types of diagonal blocks are

$$L_j \equiv \begin{bmatrix} -\lambda & 1 & & \\ & \cdot & \cdot & \\ & & -\lambda & 1 \end{bmatrix} \quad \text{and} \quad L_j^T \equiv \begin{bmatrix} -\lambda & & \\ 1 & \cdot & \\ & \cdot & -\lambda \\ & & 1 \end{bmatrix} \tag{3.3}$$

The $j \times (j+1)$ block L_j is called a *singular block of right (or column) minimal index* j. It has a one dimensional right null space, $[1, \lambda, \ldots, \lambda^j]^T$, for any λ. The $(j+1) \times j$ block L_j^T is a *singular block of left (or row) minimal index* j, and has a one dimensional left null space for any λ. The left and right singular blocks together constitute the *singular structure* of the pencil, and appear in the KCF if and only if the pencil is singular. The regular and singular structures define the *Kronecker structure* of a singular pencil.

We also have a real KCF associated with real matrix pencils. If $A, B \in \mathbf{R}^{m \times n}$, there exist nonsingular $P \in \mathbf{R}^{m \times m}$ and $Q \in \mathbf{R}^{n \times n}$, where as before $P^{-1}(A - \lambda B)Q = S - \lambda T$ is block diagonal. The only difference with (3.1) is the Jordan blocks associated with complex conjugate pairs of eigenvalues. Let $\alpha = \mu + i\omega$, where μ, ω are real and $\omega \neq 0$. If α is an eigenvalue of $A - \lambda B$, then also $\bar{\alpha}$ is an eigenvalue. Let $J_j(\alpha, \bar{\alpha})$ denote a Jordan block of size $2j \times 2j$ associated with a complex conjugate pair of eigenvalues, here illustrated with the case $j = 3$:

$$J_3(\alpha, \bar{\alpha}) \equiv \begin{bmatrix} \mu - \lambda & \omega & 1 & 0 & 0 & 0 \\ -\omega & \mu - \lambda & 0 & 1 & 0 & 0 \\ 0 & 0 & \mu - \lambda & \omega & 1 & 0 \\ 0 & 0 & -\omega & \mu - \lambda & 0 & 1 \\ 0 & 0 & 0 & 0 & \mu - \lambda & \omega \\ 0 & 0 & 0 & 0 & -\omega & \mu - \lambda \end{bmatrix}. \tag{3.4}$$

The Jordan block $J_j(\alpha, \bar{\alpha})$ plays the same role in the real Jordan canonical form as diag$(J_j(\alpha), J_j(\bar{\alpha}))$ does in the complex JCF. Notice that each pair of the $2j$ columns of the real P and Q associated with a $J_j(\alpha, \bar{\alpha})$ block form the real and imaginary parts of the (generalized) principal chains corresponding to the complex conjugate pair of eigenvalues.

3.2 Generalized Schur form and reducing subspaces

In most applications it is enough to transfer $A - \lambda B$ to a *generalized Schur form* (e.g. to GUPTRI form [11, 12])

$$P^H(A - \lambda B)Q = \begin{bmatrix} A_r - \lambda B_r & * & * \\ 0 & A_{reg} - \lambda B_{reg} & * \\ 0 & 0 & A_l - \lambda B_l \end{bmatrix}, \tag{3.5}$$

where P $(m \times m)$ and Q $(n \times n)$ are unitary and $*$ denotes arbitrary conforming submatrices. Here the square upper triangular block $A_{reg} - \lambda B_{reg}$ is regular and has the same regular structure as $A - \lambda B$ (i.e., contains all eigenvalues (finite and infinite) of $A - \lambda B$). The rectangular blocks $A_r - \lambda B_r$ and $A_l - \lambda B_l$ contain the singular structure (right and left minimal indices) of the pencil and are block upper triangular.

$A_r - \lambda B_r$ has only right minimal indices in its Kronecker canonical form (KCF), indeed the same L_j blocks as $A - \lambda B$. Similarly, $A_l - \lambda B_l$ has only left minimal indices in its KCF, the same L_j^T blocks as $A - \lambda B$. If $A - \lambda B$ is singular at least one of $A_r - \lambda B_r$ and $A_l - \lambda B_l$ will be present in (3.5). The explicit structure of the diagonal blocks in staircase form can be found in [12]. If $A - \lambda B$ is regular $A_r - \lambda B_r$ and $A_l - \lambda B_l$ are not present in (3.5) and the GUPTRI form reduces to the upper triangular block $A_{reg} - \lambda B_{reg}$. Staircase forms that reveal the Jordan structure of the zero and infinite eigenvalues are contained in $A_{reg} - \lambda B_{reg}$.

Given $A - \lambda B$ in GUPTRI form we also know different pairs of reducing subspaces [31, 11]. Suppose the eigenvalues on the diagonal of $A_{reg} - \lambda B_{reg}$ are ordered so that the first k, say, are in Λ_1 (a subset of the spectrum) and the remainder are outside Λ_1. Let $A_r - \lambda B_r$ be $m_r \times n_r$. Then the left and right reducing subspaces corresponding to Λ_1 are spanned by the leading $m_r + k$ columns of P and leading $n_r + k$ columns of Q, respectively. When Λ_1 is empty, the corresponding reducing subspaces are called *minimal*, and when Λ_1 contains the whole spectrum the reducing subspaces are called *maximal*.

Several authors have proposed (staircase-type) algorithms for computing a generalized Schur form (e.g. see [2, 21, 23, 22, 29, 34]). They are numerically stable in the sense that they compute the exact Kronecker structure (generalized Schur form or something similar) of a nearby pencil $A' - \lambda B'$. $\delta \equiv \|(A - A', B - B')\|_E$ is an upper bound on the distance to the closest $(A + \delta A, B + \delta B)$ with the KCF of (A', B'). Recently, robust software with error bounds for computing the GUPTRI form of a singular $A - \lambda B$ has been published [11, 12]. Some computational experiments that use this software will be discussed later.

3.3 Generic and non-generic Kronecker structures

Although, the KCF looks quite complicated in the general case, most matrix pencils have a quite simple Kronecker structure. 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. This corresponds to the *generic case* when $A - \lambda B$ has full rank for any complex (or real) value of λ. Accordingly, generic rectangular pencils have no regular part. The generic Kronecker structure for $A - \lambda B$ with $d = n - m > 0$ is

$$\mathrm{diag}(L_\alpha, \ldots, L_\alpha, L_{\alpha+1}, \ldots, L_{\alpha+1}), \tag{3.6}$$

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) [29, 8]. The same statement holds for $d = m - n > 0$ if we replace $L_\alpha, L_{\alpha+1}$ in (3.3) by $L_\alpha^T, L_{\alpha+1}^T$. 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 [32]:

$$\mathrm{diag}(L_j, L_{n-j-1}^T), \quad j = 0, \ldots, n-1. \tag{3.7}$$

Only if a singular $A - \lambda B$ is rank deficient (for some λ) may the associated KCF be more complicated and possibly include a regular part, as well as, right and left singular blocks. This situation corresponds to the *non-generic case*, which of course is the real challenge from a computational point of view.

The generic and non-generic cases can easily be couched in terms of reducing subspaces. For example, generic rectangular pencils have only trivial reducing subspaces and no generalized eigenvalues at all. Generic square singular pencils have the same minimal and maximal reducing subspaces. A non-generic case corresponds to that $A - \lambda B$ lies in a particular manifold of the matrix pencil space and that the pencil has nontrivial reducing subspaces. Moreover, only if it is perturbed so as to move continuously within that manifold do its reducing subspaces and generalized eigenvalues also move continuously and satisfy interesting error bounds [9, 11]. These requirements are natural in many control and systems theoretic problems such as computing controllable subspaces and uncontrollable modes.

4 The Geometry of Matrix Pencil Space

In this section we derive formulas for the tangent and normal spaces of the orbit of a matrix pencil that we will make use of in order to compute the versal form in the next section. We also derive new bounds for the distance to less generic pencils.

4.1 The orbit of a matrix pencil and its tangent and normal spaces

Any $m \times n$ matrix pair (A, B) (with real or complex entries) defines a manifold of *strictly equivalent* matrix pencils in the $2mn$ dimensional space $\mathcal{P}$ of m-by-n pencils:

$$\mathrm{orbit}(A - \lambda B) = \{P^{-1}(A - \lambda B)Q : \det(P)\det(Q) \neq 0\}. \tag{4.1}$$

We may choose a special element of $\mathrm{orbit}(A - \lambda B)$ that reveals the KCF of the pencil.

As usual the dimension of $\mathrm{orbit}(A - \lambda B)$ is equal to the dimension of the tangent space to the orbit at $A - \lambda B$, here denoted $\tan(A - \lambda B)$. By considering the deformation $(I_m + \delta X)(A - \lambda B)(I_n - \delta Y)$ of $A - \lambda B$ to first order term in δ, where δ is a small scalar, we obtain $A - \lambda B + \delta(X(A - \lambda B) - (A - \lambda B)Y) + O(\delta^2)$, from which it is evident that $\tan(A - \lambda B)$ consists of the pencils that can be represented in the form

$$T_A - \lambda T_B = (XA - AY) - \lambda(XB - BY), \tag{4.2}$$

where X is an $m \times m$ matrix and Y is an $n \times n$ matrix.

Using Kronecker products we can represent the $2mn$-vectors $T_A - \lambda T_B \in \tan(A - \lambda B)$ as

$$\begin{bmatrix} \mathrm{vec}(T_A) \\ \mathrm{vec}(T_B) \end{bmatrix} = \begin{bmatrix} A^T \otimes I_m \\ B^T \otimes I_m \end{bmatrix} \mathrm{vec}(X) - \begin{bmatrix} I_n \otimes A \\ I_n \otimes B \end{bmatrix} \mathrm{vec}(Y).$$

In this notation, we may say that the tangent space is the range of the $2mn \times (m^2 + n^2)$ matrix

$$T \equiv \begin{bmatrix} A^T \otimes I_m & -I_n \otimes A \\ B^T \otimes I_m & -I_n \otimes B \end{bmatrix}. \tag{4.3}$$

We may define the *normal* space, $\mathrm{nor}(A - \lambda B)$, as the space perpendicular to $\tan(A - \lambda B)$. Orthogonality in $\mathcal{P}$, the $2mn$ dimensional space of matrix pencils is defined with respect to a Frobenius inner product

$$< A - \lambda B, C - \lambda D > \equiv \mathrm{tr}(AC^H + BD^H), \tag{4.4}$$

where $\mathrm{tr}(X)$ denotes the trace of a square matrix X. Remembering that the space orthogonal to the range of a matrix is the kernel of the Hermitian transpose, we have that

$$\mathrm{nor}(A - \lambda B) = \ker(T^H) = \ker \begin{bmatrix} \bar{A} \otimes I_m & \bar{B} \otimes I_m \\ -I_n \otimes A^H & -I_n \otimes B^H \end{bmatrix}.$$

In ordinary matrix notation, this states that $Z_A - \lambda Z_B$ is in the normal space of $A - \lambda B$ if and only if

$$Z_A A^H + Z_B B^H = 0 \quad \text{and} \quad A^H Z_A + B^H Z_B = 0. \tag{4.5}$$

The conditions on Z_A and Z_B can easily be verified and also be derived in terms of the Frobenius inner product, i.e.,

$$< T_A - \lambda T_B, Z_A - \lambda Z_B > = \operatorname{tr}(X(AZ_A^H + BZ_B^H) - (Z_A^H A + Z_B^H B)Y). \tag{4.6}$$

Verification: if the conditions (4.5) are satisfied, it follows from (4.6) that the inner product is zero. Derivation: if $< T_A - \lambda T_B, Z_A - \lambda Z_B > = 0$, then $\operatorname{tr}(X(AZ_A^H + BZ_B^H) - (Z_A^H A + Z_B^H B)Y) = 0$ must hold for any X (of size $m \times m$) and Y (of size $n \times n$). By choosing $X \equiv 0$, (4.6) reduces to $\operatorname{tr}((Z_A^H A + Z_B^H B)Y) = 0$, which holds for any Y if and only if $Z_A^H A + Z_B^H B = 0$. Similarly, we can chose $Y \equiv 0$, which gives that $AZ_A^H + BZ_B^H = 0$.

If $B = I$, this reduces to $Z_A \in \operatorname{nor}(\mathrm{A})$ if and only if $Z_A^H \in \operatorname{centralizer}(A)$, which is a well-known fact (e.g. see [1]). We will see in Section 5.3 that though the A-part of the normal space is very simple when $B = I$, obtaining an orthonormal basis for the B-part is particularly challenging. The requirement that $Z_B = -A^H Z_A$ when $B = I$ destroys any orthogonality one may have in a basis for the A-part.

We now collect our general statements and a few obvious consequences:

Theorem 4.1 *Let the $m \times n$ pencil $A - \lambda B$ be given. Define the $2mn \times (m^2 + n^2)$ matrix T as in (4.3). Then*

$$\tan(A - \lambda B) = \operatorname{range}(T) = \{(XA - AY) - \lambda(XB - BY)\},$$

where X and Y are compatible square matrices, and

$$\operatorname{nor}(A - \lambda B) = \ker(T^H) = \{Z_A - \lambda Z_B\},$$

where $Z_A A^H + Z_B B^H = 0$ and $\quad A^H Z_A + B^H Z_B = 0$.
The dimensions of these spaces are

$$\dim(\tan(A - \lambda B)) = m^2 + n^2 - \dim(\ker(T)), \tag{4.7}$$

and

$$\dim(\operatorname{nor}(A - \lambda B)) = \dim(\ker(T^H)) = \dim(\ker(T)) - (m - n)^2. \tag{4.8}$$

Of course, the tangent and normal spaces are complementary and span the complete $2mn$ dimensional space, i.e., $\mathcal{P} = \tan(A - \lambda B) \oplus \operatorname{nor}(A - \lambda B)$, so that the dimensions in (4.7) and (4.8) add up to $2mn$ as they should.

Theorem 4.1 leads to one approach for computing a basis for $\operatorname{nor}(A - \lambda B)$ from the singular value decomposition (SVD) of T. Indeed, the left singular vectors corresponding to the zero singular value form such a basis. The dimension of the normal space is also known as the *codimension* of the orbit, here denoted $\operatorname{cod}(A - \lambda B)$. Accordingly, we have the following "compact" characterization of the codimension of $\operatorname{orbit}(A - \lambda B)$.

Corollary 4.1 *Let the $m \times n$ pencil $A - \lambda B$ be given. Then,*

$$\operatorname{cod}(A - \lambda B) = \text{the number of zero singular values of } T. \tag{4.9}$$

The corresponding result for the (square) matrix case is

$$\text{cod}(A) = \text{the number of zero singular values of } I_n \otimes A - A^T \otimes I_n.$$

Although, the *SVD*-based method is simple and has nice numerical properties (backward stability), it is rather costly in the number of operations. Computing the *SVD* of T is an $O(m^3n^3)$ operation.

Knowing the Kronecker structure of $A - \lambda B$, it is also possible to compute the codimension of the orbit as the sum of separate codimensions [8]:

$$\text{cod}(A - \lambda B) = c_{\text{Jor}} + c_{\text{Right}} + c_{\text{Left}} + c_{\text{Jor,Sing}} + c_{\text{Sing}}. \tag{4.10}$$

The different contributions in (4.10) originate from the Jordan structure of all eigenvalues (including any infinite eigenvalue), the right singular blocks ($L_j \leftrightarrow L_k$), the left singular blocks ($L_j^T \leftrightarrow L_k^T$), interactions of the Jordan structure with the singular blocks (L_k and L_j^T) and interactions between the left and right singular structures ($L_j \leftrightarrow L_k^T$), respectively. Explicit expressions for these codimensions are derived in [8]. Assume that the given $A - \lambda B$ has $p \leq \min(m, n)$ distinct eigenvalues $\lambda_i, i = 1 : p$ with p_i Jordan blocks each. Let $q_1(\lambda_i) \geq q_2(\lambda_i) \geq \ldots \geq q_{p_i}(\lambda_i)$ denote the sizes of the Jordan blocks corresponding to the eigenvalue λ_i. Then the separate codimensions of (4.10) can be expressed as

$$c_{\text{Jor}} = \sum_{i=1}^{p}\sum_{j=1}^{p_i}(2j-1)q_j(\lambda_i) = \sum_{i=1}^{p}(q_1(\lambda_i) + 3q_2(\lambda_i) + 5q_3(\lambda_i) + \ldots),$$

$$c_{\text{Right}} = \sum_{j>k}(j-k-1), \quad c_{\text{Left}} = \sum_{j>k}(j-k-1), \quad c_{\text{Sing}} = \sum_{j,k}(j+k+2),$$

$$c_{\text{Jor,Sing}} = (\text{size of complete regular part}) \cdot (\text{number of singular blocks}).$$

Notice that if we do not wish to specify the value of an eigenvalue λ_i, the codimension count for this unspecified eigenvalue is one less, i.e.,

$$-1 + q_1(\lambda_i) + 3q_2(\lambda_i) + 5q_3(\lambda_i) + \ldots$$

This is sometimes done in algorithms for computing the Kronecker structure of a matrix pencil, where usually only the eigenvalues 0 and ∞ are specified and the remaining ones are unspecified.

It is possible to extract the Kronecker structure of $A - \lambda B$ from a generalized Schur decomposition in $O((\max(m, n))^3)$ operations. The most reliable *SVD*-approach for computing a generalized Schur decomposition of $A - \lambda B$ requires at most $O((\max(m, n))^4)$ operations, which is still small compared to computing the *SVD* of T (4.3) for already moderate values of m and n (e.g. when $m = n$).

For given m and n the generic pencil has codimension 0 (i.e., span the complete $2mn$ dimensional space) while the most non-generic matrix pair $(A, B) = (0_{m\times n}, 0_{m\times n})$ has codimension $= 2mn$ (i.e., defines a "point" in $2mn$ dimensional space). Accordingly, any $m \times n$ non-generic pencil different from the "zero pencil" has a codimension ≥ 1 and $< 2mn$.

4.2 A lower bound on the distance to a less generic pencil

The *SVD* characterization of the codimension of orbit$(A-\lambda B)$ in Corollary 4.1 leads to the following theorem from which we present an interesting special case as a corollary.

Theorem 4.2 *For a given $m \times n$ pencil $A-\lambda B$ with codimension c, a lower bound on the distance to the closest pencil $(A+\delta A)-\lambda(B+\delta B)$ with codimension $c+d$, where $d \geq 1$ is given by*

$$\|(\delta A, \delta B)\|_E \geq \frac{1}{\sqrt{m+n}} \left(\sum_{i=2mn-c-d+1}^{2mn} \sigma_i^2(T) \right)^{1/2}, \tag{4.11}$$

where $\sigma_i(T)$ denotes the ith largest singular value of T ($\sigma_i(T) \geq \sigma_{i+1}(T) \geq 0$).

Proof It follows from Corollary 4.1 that T has rank $= 2mn-c$ if and only if $A-\lambda B$ has codimension c and $(A+\delta A)-\lambda(B+\delta B)$ has codimension $c+d$, $(d \geq 1)$ if and only if $T+\delta T$, where δT is defined as

$$\delta T \equiv \begin{bmatrix} \delta A^T \otimes I_m & -I_n \otimes \delta A \\ \delta B^T \otimes I_m & -I_n \otimes \delta B \end{bmatrix}, \tag{4.12}$$

has rank $2mn-c-d$. From the construction, it follows that $\|\delta T\|_E = \sqrt{m+n}\|(\delta A, \delta B)\|_E$ (each element δa_{ij} and δb_{ij} appears $m+n$ times in δT). The Eckart–Young and Mirsky theorem for finding the closest matrix of a given rank (e.g. see [16]), gives that the size of the smallest perturbation in Frobenius norm that reduces the rank in T from $2mn-c$ to $2mn-c-d$ is

$$\left(\sum_{i=2mn-c-d+1}^{2mn-c} \sigma_i^2(T) \right)^{1/2}. \tag{4.13}$$

Moreover, $A-\lambda B$ has codimension c implies that $\sigma_{2mn-c+1}(T) = \ldots = \sigma_{2mn}(T) = 0$. Since $\|\delta T\|_E$ must be larger than or equal to the quantity (4.13), the proof is complete. □

Corollary 4.2 *For a given generic $m \times n$ pencil $A-\lambda B$, a lower bound on the distance to the closest non-generic pencil $(A+\delta A)-\lambda(B+\delta B)$ is given by*

$$\|(\delta A, \delta B)\|_E \geq \frac{\sigma_{\min}(T)}{\sqrt{m+n}}, \tag{4.14}$$

where $\sigma_{\min}(T) = \sigma_{2mn}(T)$ denotes the smallest singular value of T, which is non-zero for a generic $A-\lambda B$.

We remark that the set of $m \times n$ matrix pencils does not include orbits of all codimensions from 1 to $2mn$.

One application of Corollary 4.2 is to characterize the distance to uncontrollability for a multiple input multiple output linear system $E\dot{x}(t) = Fx(t) + Gu(t)$, where E and F are p-by-p matrices, G is p-by-q ($p \geq q$), and E is assumed to be nonsingular. If $A-\lambda B \equiv [G|F-\lambda E]$ is generic, the linear system is controllable (i.e., the dimension of the controllable subspace equals p) and a lower bound on the distance to the closest uncontrollable system is given by (4.14).

5 Versal Deformations for the Kronecker Canonical Form

In this section, we derive versal deformations which for us will mean the decomposition of arbitrary perturbations into the tangent and normal spaces of the orbits of equivalent pencils.

5.1 An introductory example

We start with a small example before considering the general case. Let $A - \lambda B = L_1 \oplus L_4$ with codimension $= 2$. (This means that the manifold orbit$(A - \lambda B)$ has codimension 2 or dimension 68 in the 70 dimensional space of 5×7 pencils.) Since $A - \lambda B$ already is in KCF we know its block structure:

$$A - \lambda B = \left[\begin{array}{cc|ccccc} -\lambda & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & -\lambda & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\lambda & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\lambda & 1 \end{array}\right]. \tag{5.1}$$

From (4.2) the matrices in the tangent space are given by $T_A - \lambda T_B = (XA - AX) - \lambda(XB - BX)$, where

$$T_A = \left[\begin{array}{cc|ccccc} -y_{21} & x_{11} - y_{22} & -y_{23} & x_{12} - y_{24} & x_{13} - y_{25} & x_{14} - y_{26} & x_{15} - y_{27} \\ \hline \boxed{\boxed{-y_{41}}} & x_{21} - y_{42} & -y_{43} & x_{22} - y_{44} & x_{23} - y_{45} & x_{24} - y_{46} & x_{25} - y_{47} \\ \boxed{-y_{51}} & \boxed{\boxed{x_{31} - y_{52}}} & -y_{53} & x_{32} - y_{54} & x_{33} - y_{55} & x_{34} - y_{56} & x_{35} - y_{57} \\ -y_{61} & \boxed{x_{41} - y_{62}} & -y_{63} & x_{42} - y_{64} & x_{43} - y_{65} & x_{44} - y_{66} & x_{45} - y_{67} \\ -y_{71} & x_{51} - y_{72} & -y_{73} & x_{52} - y_{74} & x_{53} - y_{75} & x_{54} - y_{76} & x_{55} - y_{77} \end{array}\right] \tag{5.2}$$

and

$$T_B = \left[\begin{array}{cc|ccccc} x_{11} - y_{11} & -y_{12} & x_{12} - y_{13} & x_{13} - y_{14} & x_{14} - y_{15} & x_{15} - y_{16} & -y_{17} \\ \hline x_{21} - y_{31} & -y_{32} & x_{22} - y_{33} & x_{23} - y_{34} & x_{24} - y_{35} & x_{25} - y_{36} & -y_{37} \\ \boxed{\boxed{x_{31} - y_{41}}} & -y_{42} & x_{32} - y_{43} & x_{33} - y_{44} & x_{34} - y_{45} & x_{35} - y_{46} & -y_{47} \\ \boxed{x_{41} - y_{51}} & \boxed{\boxed{-y_{52}}} & x_{42} - y_{53} & x_{43} - y_{54} & x_{44} - y_{55} & x_{45} - y_{56} & -y_{57} \\ x_{51} - y_{61} & \boxed{-y_{62}} & x_{52} - y_{63} & x_{53} - y_{64} & x_{54} - y_{65} & x_{55} - y_{66} & -y_{67} \end{array}\right]. \tag{5.3}$$

By inspection we find the following two relations between elements in T_A and T_B:

$$\boxed{\boxed{}}: \; t^a_{21} + t^a_{32} = t^b_{31} + t^b_{42}, \tag{5.4}$$

and

$$\boxed{}: \; t^a_{31} + t^a_{42} = t^b_{41} + t^b_{52} \tag{5.5}$$

where t^a_{ij} and t^b_{ij} denote the (i, j)-th elements of T_A and T_B, respectively. These two relations show clearly that the tangent space has codimension at least two. It may be verified that the other parameters may be chosen arbitrarily so that the codimension is exactly two.

We want to find $Z_A - Z_B$ that is orthogonal to $T_A - \lambda T_B$ with respect to the Frobenius inner product, i.e.,

$$0 \equiv \; < T_A - \lambda T_B, Z_A - \lambda Z_B > \; \equiv \mathrm{tr}(T_A Z_A^H + T_B Z_B^H) \equiv \sum_{i,j} t_{ij}^a \overline{z}_{ij}^a + t_{ij}^b \overline{z}_{ij}^b. \tag{5.6}$$

This inner product is most easily envisioned as the sum of the elementwise multiplication of the two pencils. Using this point of view it is obvious that the normal space consists of pencils of the form $Z_A - \lambda Z_B \in \mathrm{nor(A - \lambda B)}$:

$$\begin{aligned} Z_A - \lambda Z_B \;\; &= \left[\begin{array}{cc|ccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline p_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_2 & p_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & p_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] - \lambda \left[\begin{array}{cc|ccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -p_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -p_2 & -p_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -p_2 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \\ &= \left[\begin{array}{cc|ccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline p_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_2 + \lambda p_1 & p_1 & 0 & 0 & 0 & 0 & 0 \\ \lambda p_2 & p_2 + \lambda p_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda p_2 & 0 & 0 & 0 & 0 & 0 \end{array}\right], \end{aligned} \tag{5.7}$$

where p_1 and p_2 are arbitrary. Roughly speaking the parameter p_1 corresponds to the doubly boxed entries ($\boxed{\square}$) and the parameter p_2 corresponds to the singly boxed entries. ($\boxed{}$).

Now, $A - \lambda B + Z_A - \lambda Z_B$ may be thought of as a versal deformation, or normal form, with minimum number of parameters (equal to the codimension of the original pencil). It follows that any (complex) pencil close to the given $A - \lambda B$ can be reduced to the 2-parameter normal form $A - \lambda B + Z_A - \lambda Z_B$, where $A - \lambda B$ is in Kronecker canonical form.

5.2 Notation: a glossary of Toeplitz and Hankel matrices

The example in the previous section shows that a non-zero block of $Z_A - \lambda Z_B$ has a structured form. Indeed, the $(2,1)$ block has a Toeplitz-like form with $j - i = 3$ non-zero diagonals starting from the $(1,1)$-element of the $(2,1)$ block. A closer look shows that the A-part has $i - j - 1 = 2$ non-zero diagonals and the B-part is just the same matrix negated and with the diagonals shifted one row downwards. In general, different non-zero blocks with Toeplitz or Hankel properties will show up in $Z_A - \lambda Z_B \in \mathrm{nor}(A - \lambda B)$. To simplify the proof of the general case we introduce some Toeplitz and Hankel matrices. Arrows and "stops" near the matrices make clear how the matrix is defined.

Let $S_{s \times t}^L$ be a *lower trapezoidal* s-by-t Toeplitz matrix with the first non-zero diagonal starting at position $(1,1)$:

$$S^L_{s\times t} = \begin{array}{c} \downarrow \\ \\ \\ \\ \\ \\ \perp \end{array}\begin{bmatrix} p_1 & 0 & 0 \\ \vdots & \ddots & 0 \\ \vdots & & p_1 \\ p_{s-t+1} & & \vdots \\ \vdots & \ddots & \vdots \\ p_s & \cdots & p_{s-t+1} \end{bmatrix} \text{ if } s \geq t, \text{ and } S^L_{s\times t} = \begin{array}{c} \downarrow \\ \\ \perp \end{array}\begin{bmatrix} p_1 & 0 & & \cdots & & 0 \\ \vdots & \ddots & \ddots & & & \vdots \\ p_s & \cdots & p_1 & 0 & \cdots & 0 \end{bmatrix}, \text{ otherwise,}$$

and let $T^L_{s\times t}$ be a *lower trapezoidal* s-by-t Toeplitz matrix with the first non-zero diagonal's last element at position (s,t):

$$T^L_{s\times t} = \begin{array}{c} \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & & \vdots \\ p_1 & \ddots & \vdots \\ \vdots & \ddots & 0 \\ p_t & \cdots & p_1 \end{bmatrix} \\ \begin{array}{ccc} \vdash & \quad & \leftarrow \end{array} \end{array} \text{ if } s \geq t, \text{ and } T^L_{s\times t} = \begin{array}{c} \begin{bmatrix} p_{t-s+1} & \cdots & p_1 & 0 & 0 \\ \vdots & \ddots & & \ddots & 0 \\ p_t & \cdots & p_{t-s+1} & \cdots & p_1 \end{bmatrix} \\ \begin{array}{ccccc} \vdash & \quad & \quad & \quad & \leftarrow \end{array} \end{array}, \text{ otherwise.}$$

If $s < t$, the entries of the last $t-s$ columns of $S^L_{s\times t}$ are zero. Similarly, if $s \geq t$, the entries of the first $s-t$ rows of $T^L_{s\times t}$ are zero.

Let $S^B_{s\times t}$ be a *banded lower trapezoidal* s-by-t Toeplitz with last row 0:

$$S^B_{s\times t} = \begin{array}{c} \downarrow \\ \\ \\ \perp \\ \\ \\ \\ \end{array}\begin{bmatrix} p_1 & 0 & 0 \\ \vdots & \ddots & 0 \\ \vdots & & p_1 \\ p_{s-t} & & \vdots \\ 0 & \ddots & \vdots \\ \vdots & \ddots & p_{s-t} \\ 0 & \cdots & 0 \end{bmatrix} \text{ if } s > t, \text{ and } S^B_{s\times t} = 0, \text{ otherwise,}$$

and let $T^B_{s\times t}$ be another *banded lower trapezoidal* s-by-t Toeplitz matrix, this time with last column 0:

$$T^B_{s\times t} = \begin{array}{c} \begin{array}{cccccc} \vdash & \quad & \quad & \leftarrow & \quad & \quad \end{array} \\ \begin{bmatrix} p_{t-s} & \cdots & p_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \ddots & \ddots & \vdots \\ 0 & 0 & p_{t-s} & \cdots & p_1 & 0 \end{bmatrix} \end{array} \text{ if } s < t, \text{ and } T^B_{s\times t} = 0, \text{ otherwise.}$$

Notice that the last row of $S^B_{s\times t}$ (if $s > t$) and the last column of $T^B_{s\times t}$ (if $s < t$) have all entries equal to zero.

Moreover, let $H^L_{s\times t}$ be a *lower trapezoidal* s-by-t Hankel matrix with the first non-zero diagonal starting at position $(1, t)$:

$$
H^L_{s\times t} = \begin{array}{@{}c@{}c@{}} \begin{bmatrix} 0 & 0 & p_1 \\ 0 & \iddots & \vdots \\ p_1 & & \vdots \\ \vdots & & p_{s-t+1} \\ \vdots & \iddots & \vdots \\ p_{s-t+1} & \cdots & p_s \end{bmatrix} & \begin{array}{c} \downarrow \\ \\ \\ \\ \\ \perp \end{array} \end{array} \quad \text{if } s \geq t, \text{ and } \quad H^L_{s\times t} = \begin{array}{@{}c@{}c@{}} \begin{bmatrix} 0 & \cdots & & 0 & p_1 \\ \vdots & & \iddots & \iddots & \vdots \\ 0 \cdots 0 & p_1 & & \cdots & p_s \end{bmatrix} & \begin{array}{c} \downarrow \\ \\ \perp \end{array} \end{array}, \text{ otherwise,}
$$

and let $H^U_{s\times t}$ be a similar *upper trapezoidal* s-by-t Hankel matrix:

$$
H^U_{s\times t} = \begin{array}{c} \begin{array}{ccc} \vdash & & \leftarrow \end{array} \\ \begin{bmatrix} p_t & \cdots & p_1 \\ \vdots & \iddots & 0 \\ p_1 & \iddots & \vdots \\ 0 & & \vdots \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \end{array} \quad \text{if } s \geq t, \text{ and } \quad H^U_{s\times t} = \begin{array}{c} \begin{array}{ccccc} \vdash & & & & \leftarrow \end{array} \\ \begin{bmatrix} p_t & \cdots & p_{t-s+1} & \cdots & p_1 \\ \vdots & \iddots & & \iddots & 0 \\ p_{t-s+1} & \cdots & p_1 & 0 & 0 \end{bmatrix} \end{array}, \text{ otherwise.}
$$

If $s < t$, the entries of the first $t - s$ columns of $H^L_{s\times t}$ are zero. Similarly, if $s \geq t$, the entries of the last $s - t$ rows of $H^U_{s\times t}$ are zero.

Let $H_{s\times t}$ be a *dense* s-by-t Hankel matrix (with the first diagonal starting at position $(1, 1)$):

$$
H_{s\times t} = \begin{bmatrix} p_1 & p_2 & p_3 \cdots & p_t \\ p_2 & \iddots & & \vdots \\ p_3 & & & \vdots \\ \vdots & & & \vdots \\ p_s & & \cdots & p_{s+t-1} \end{bmatrix},
$$

for both the cases $s \geq t$ and $s < t$.

The *nilpotent* k-by-k matrix

$$
C_k = \begin{bmatrix} 0 & I_{k-1} \\ 0 & 0 \end{bmatrix},
$$

will be used as a shift operator. For a given k-by-n matrix X, the rows are shifted one row upwards and downwards by the operations $C_k X$ and $C_k^T X$, respectively. The columns are shifted one column rightwards and one column leftwards in an n-by-k matrix X by the operations XC_k and XC_k^T, respectively. The k-by-$(k + 1)$ matrices

$$
G_k = [I_k \; 0] \text{ and } \hat{G}_k = [0 \; I_k],
$$

will be used to pick all rows but one or all columns but one of a given matrix X in the following way. The first k and last k rows in a $(k + 1)$-by-n matrix X are picked by $G_k X$

and $\hat{G}_k X$, respectively. The k first and k last columns in an n-by-$(k+1)$ matrix X are picked by XG_k^T and $X\hat{G}_k^T$, respectively.

Let $\hat{I}_k$ denote the k-by-k matrix obtained by reversing the order of the columns in the k-by-k identity matrix. It follows that for an n-by-k matrix X, the order of the columns is reversed by the multiplication $X\hat{I}_k$.

So far, the matrices introduced are rectangular Toeplitz and Hankel matrices with a special structure, e.g. lower trapezoidal (S^L, T^L, H^L), banded lower trapezoidal (S^B, T^B), upper trapezoidal (H^U) or dense (H). The matrices C and $G, \hat{G}$ that will be used as "shift" and "pick" operators, respectively, are Toeplitz matrices with only one non-zero diagonal. In the next section we will see that versal deformations for all combinations of different blocks in the KCF, except Jordan blocks with non-zero, finite eigenvalues, can be expressed in terms of these matrices. To cope with non-zero, finite Jordan blocks $J_k(\gamma), \gamma \neq 0$ we need to introduce three more matrices. First, two lower triangular Toeplitz matrices D^L and E^L which are involved in the case with two $J_k(\gamma)$ blocks. Finally, the "monstrous" matrix F^D, which captures the cases with a (left or right) singular block and a $J_k(\gamma)$ block.

Given $\gamma \neq \{0, \infty\}$, define two infinite sequences of numbers d_i and e_i by the recursion

$$\begin{bmatrix} d_i \\ \gamma e_i \end{bmatrix} = -\begin{bmatrix} 1 & 1 \\ 1 & 2-1/i \end{bmatrix} \begin{bmatrix} \overline{\gamma} d_{i-1} \\ e_{i-1} \end{bmatrix}, \tag{5.8}$$

starting with

$$\begin{bmatrix} d_1 \\ e_1 \end{bmatrix} = \begin{bmatrix} \gamma \\ 1 \end{bmatrix}.$$

Given sizes s and t, for $1 \leq q \leq \min\{s,t\}$, we define $D_{s\times t}[q]$ and $E_{s\times t}[q]$ as *lower triangular* Toeplitz matrices with q diagonals in terms of $d_1, \ldots, d_q$ and $e_1, \ldots, e_{q-1}$ and a *boundary* value $e_q^* = -\overline{\gamma} d_q$.

$$D_{s\times t}[q] = \begin{bmatrix} 0 & & \cdots & & 0 \\ d_q & & & & \\ d_{q-1} & \ddots & & & \\ \vdots & \ddots & \ddots & & \vdots \\ d_2 & & \ddots & \ddots & \\ d_1 & d_2 & \cdots & d_{q-1} \; d_q & 0 \end{bmatrix} \quad \text{and} \quad E_{s\times t}[q] = \begin{bmatrix} 0 & & \cdots & & 0 \\ e_q^* & & & & \\ e_{q-1} & \ddots & & & \\ \vdots & \ddots & \ddots & & \vdots \\ e_2 & & \ddots & \ddots & \\ e_1 & e_2 & \cdots & e_{q-1} \; e_q^* & 0 \end{bmatrix}.$$

We take linear combinations with parameters p_j to form the matrices

$$D_{s\times t}^L = \sum_{i=1}^{\min\{s,t\}} p_j D_{s\times t}[i]\, \pi(i), \quad \text{and} \quad E_{s\times t}^L = \sum_{i=1}^{\min\{s,t\}} p_j E_{s\times t}[i]\, \pi(i), \tag{5.9}$$

where $j = \min\{s,t\} - i + 1$ and $\pi(i) = -\prod_{k=2}^{i-1} k\gamma/(1-2k)$ is defined to be $1/\gamma$ and -1 for $i = 1$ and $i = 2$, respectively. The parameter index j and the scaling function $\pi(i)$ are chosen to satisfy $D_{s\times t}^L = S_{s\times t}^L$ and $E_{s\times t}^L = -C_s^T S_{s\times t}^L$ for $\gamma = 0$ in Theorem 5.1 (see tables 1 and 2). By simplifying (5.9) using $i = j$ and $\pi(i) = 1$ this consistency will be lost, but we will still have valid expressions for the versal deformations.

The relations between the elements of $D^L_{s\times t}$ and $E^L_{s\times t}$ are most readily shown by an example:

$$D^L_{4,3} = \begin{bmatrix} 0 & 0 & 0 \\ p_1\left(\frac{2|\gamma|^4}{3}+\frac{4|\gamma|^2}{3}+1\right) & 0 & 0 \\ p_1\left(-\frac{2\gamma|\gamma|^2}{3}-\frac{2\gamma}{3}\right)+p_2\left(|\gamma|^2+1\right) & p_1\left(\frac{2|\gamma|^4}{3}+\frac{4|\gamma|^2}{3}+1\right) & 0 \\ p_1\frac{2\gamma^2}{3}-p_2\gamma+p_3 & p_1\left(-\frac{2\gamma|\gamma|^2}{3}-\frac{2\gamma}{3}\right)+p_2\left(|\gamma|^2+1\right) & p_1\left(\frac{2|\gamma|^4}{3}+\frac{4|\gamma|^2}{3}+1\right) \end{bmatrix},$$

and

$$E^L_{4,3} = \begin{bmatrix} 0 & 0 & 0 \\ p_1\left(-\frac{2\overline{\gamma}|\gamma|^4}{3}-\frac{4\overline{\gamma}|\gamma|^2}{3}-\overline{\gamma}\right) & 0 & 0 \\ p_1\left(-\frac{2|\gamma|^2}{3}-1\right)+p_2\left(-\overline{\gamma}|\gamma|^2-\overline{\gamma}\right) & p_1\left(-\frac{2\overline{\gamma}|\gamma|^4}{3}-\frac{4\overline{\gamma}|\gamma|^2}{3}-\overline{\gamma}\right) & 0 \\ p_1\frac{2\gamma}{3}-p_2-p_3\overline{\gamma} & p_1\left(-\frac{2|\gamma|^2}{3}-1\right)+p_2\left(-\overline{\gamma}|\gamma|^2-\overline{\gamma}\right) & p_1\left(-\frac{2\overline{\gamma}|\gamma|^4}{3}-\frac{4\overline{\gamma}|\gamma|^2}{3}-\overline{\gamma}\right) \end{bmatrix}.$$

Let $F^D_{s\times t}$ (D for dense) be defined as

$$F^D_{s\times t} = \sum_{i=1}^{s} p_{s-i+1}\ F_{s\times t}[i],$$

where $F_{s\times t}[q]$ has the q last rows non-zero and defined as:

$$\begin{array}{ll} f_{s-q+1,j} = \overline{\gamma}^{j-1} & \text{for } j = 1,\ldots,t, \\ f_{i,j} = \overline{\gamma} f_{i,j-1} + f_{i-1,j-1} & \text{for } i = s-q+2,\ldots,s,\ j = 2,\ldots,t, \end{array} \tag{5.10}$$

and $f_{i,1}$ for $i = s-q+2,\ldots,s$ is defined as the solution to

$$< F_{s\times t}[q]G^T_{t-1} - \lambda F_{s\times t}[q]\hat{G}^T_{t-1},\ F_{s\times t}[s-i+1]G^T_{t-1} - \lambda F_{s\times t}[s-i+1]\hat{G}^T_{t-1} > \equiv 0.$$

Notice that $f_{i,1}$ is used as an unknown in the generation of elements in (5.10). In the definition of $F_{s\times t}[q]$, the solutions for $f_{i,1}$ for $i = s-q+2,\ldots,s$ ensure that $F_{s\times t}[q]G^T_{t-1} - \lambda F_{s\times t}[q]\hat{G}^T_{t-1}$ is orthogonal to $F_{s\times t}[\hat{q}]G^T_{t-1} - \lambda F_{s\times t}[\hat{q}]\hat{G}^T_{t-1}$ for $\hat{q} = 1,\ldots,q-1$.

Also here we show a small example to facilitate the interpretation of the definition:

$$F^D_{3\times 2} = \begin{bmatrix} p_1 & p_1\overline{\gamma} \\ p_2 - p_1\frac{(|\gamma|^2+1)\gamma}{|\gamma|^4+2|\gamma|^2+2} & p_2\overline{\gamma} + p_1\frac{|\gamma|^2+2}{|\gamma|^4+2|\gamma|^2+2} \\ p_3 - p_2\frac{\gamma}{|\gamma|^2+1} + p_1\frac{\gamma^2}{|\gamma|^4+2|\gamma|^2+2} & p_3\overline{\gamma} + p_2\frac{1}{|\gamma|^2+1} - p_1\frac{\gamma}{|\gamma|^4+2|\gamma|^2+2} \end{bmatrix}$$

5.3 Versal deformations – the general case

Without loss of generality assume that $A - \lambda B$ already is in Kronecker canonical form, $M = \mathrm{diag}(M_1, M_2, \ldots, M_b)$, where each M_k is either a Jordan block associated with a finite or infinite eigenvalue or a singular block corresponding to a left or right minimal index. A pencil $T_A - \lambda T_B = XM - MY$ in the tangent space can be partitioned conformally with the pencil M so that $T^A_{ij} - \lambda T^B_{ij} = X_{ij}M_j - M_iY_{ij}$, where M_k is m_k-by-n_k, X_{ij} is m_i-by-m_j and Y_{ij} is n_i-by-n_j:

$$\begin{bmatrix} X_{11} & \cdots & X_{1b} \\ \vdots & \ddots & \vdots \\ X_{b1} & \cdots & X_{bb} \end{bmatrix} \begin{bmatrix} M_1 & & \\ & \ddots & \\ & & M_b \end{bmatrix} - \begin{bmatrix} M_1 & & \\ & \ddots & \\ & & M_b \end{bmatrix} \begin{bmatrix} Y_{11} & \cdots & Y_{1b} \\ \vdots & \ddots & \vdots \\ Y_{b1} & \cdots & Y_{bb} \end{bmatrix}. \tag{5.11}$$

Since the blocks $T^A_{ij} - \lambda T^B_{ij}, i, j = 1, \ldots, b$ are mutually independent, we can study the different blocks of $T_A - \lambda T_B$ separately. Let $Z^A_{ij} - \lambda Z^B_{ij}$ be conformally sized blocks of $Z_A - \lambda Z_B$. From (4.5) we know that $Z_A - \lambda Z_B$ is in the normal space if and only if $A^H Z_A + B^H Z_B = 0$ and $Z_A A^H + Z_B B^H = 0$. We obtain a simple result since A and B are block diagonal.

Proposition 5.1 *Assume that $M = A - \lambda B = \mathrm{diag}(A_1, A_2, \ldots, A_b) - \lambda \mathrm{diag}(B_1, B_2, \ldots, B_b)$ is in Kronecker canonical form, where each block $A_i - \lambda B_i \equiv M_i$ represents one block in the Kronecker structure. Then $Z_A - \lambda Z_B \in \mathrm{nor}(A - \lambda B)$ if and only if*

$$A_j^H Z^A_{ji} = -B_j^H Z^B_{ji} \quad \text{and} \quad Z^A_{ji} A_i^H = -Z^B_{ji} B_i^H, \quad \text{for } i = 1, \ldots, b \text{ and } j = 1, \ldots, b. \tag{5.12}$$

The mutual independency of the (i, j) blocks of Z_A and Z_B implies that we only have to consider two M_k blocks at a time:

$$T_A[i,j] - \lambda T_B[i,j] = \begin{bmatrix} X_{ii} & X_{ij} \\ X_{ji} & X_{jj} \end{bmatrix} \begin{bmatrix} M_i & 0 \\ 0 & M_j \end{bmatrix} - \begin{bmatrix} M_i & 0 \\ 0 & M_j \end{bmatrix} \begin{bmatrix} Y_{ii} & Y_{ij} \\ Y_{ji} & Y_{jj} \end{bmatrix} = \begin{bmatrix} T^A_{ii} & T^A_{ij} \\ T^A_{ji} & T^A_{jj} \end{bmatrix} - \lambda \begin{bmatrix} T^B_{ii} & T^B_{ij} \\ T^B_{ji} & T^B_{jj} \end{bmatrix},$$

and

$$Z_A[i,j] - \lambda Z_B[i,j] = \begin{bmatrix} Z^A_{ii} & Z^A_{ij} \\ Z^A_{ji} & Z^A_{jj} \end{bmatrix} - \lambda \begin{bmatrix} Z^B_{ii} & Z^B_{ij} \\ Z^B_{ji} & Z^B_{jj} \end{bmatrix}. \tag{5.13}$$

Notably, by interchanging the blocks $M_i = A_i - \lambda B_i$ and $M_j = A_j - \lambda B_j$ in the KCF, we only have to interchange the corresponding blocks in $Z_A - \lambda Z_B$ accordingly. For example, if $Z_A[i,j] - \lambda Z_B[i,j]$ in (5.13) belongs to $\mathrm{nor}(\mathrm{diag}(M_i, M_j))$, then

$$\begin{bmatrix} Z^A_{jj} & Z^A_{ji} \\ Z^A_{ij} & Z^A_{ii} \end{bmatrix} - \lambda \begin{bmatrix} Z^B_{jj} & Z^B_{ji} \\ Z^B_{ij} & Z^B_{ii} \end{bmatrix} \in \mathrm{nor}(\mathrm{diag}(M_j, M_i)).$$

This implies that given two blocks M_i and M_j, it is enough to consider the case $\mathrm{diag}(M_i, M_j)$. In the following we will order the blocks in the KCF so that $Z_A - \lambda Z_B$ is block lower triangular.

Theorem 5.1 *Let $A - \lambda B = \mathrm{diag}(A_1, A_2, \ldots, A_b) - \lambda \mathrm{diag}(B_1, B_2, \ldots, B_b)$ be in KCF with the structure blocks $M_i = A_i - \lambda B_i$ ordered as follows: L_k, $J_k(0)$, $J_k(\gamma)$ (for $\gamma \neq \{0, \infty\}$), N_k, and L_k^T, where the ordering within each block-type is in increasing order of size, except for the L_k^T blocks which are ordered by decreasing order of size.*

For all i and j, let the $(i,j), (j,i)$ and $(i,i), (j,j)$ blocks of $Z_A - \lambda Z_B$ corresponding to $\mathrm{diag}(M_i, M_j)$ be built from Table 1 and Table 2, respectively.

Then $Z_A - \lambda Z_B$ gives an orthogonal basis for $\mathrm{nor}(A - \lambda B)$ with minimum number of parameters.

Table 1: Blocks in $Z_A - \lambda Z_B \in \operatorname{nor}(A - \lambda B)$, where it for $L_\alpha \oplus L_\beta$, $J_\alpha(0) \oplus J_\beta(0)$, $J_\alpha(\gamma) \oplus J_\beta(\gamma)$, and $N_\alpha \oplus N_\beta$, is assumed that $\alpha \leq \beta$. For $L_\alpha^T \oplus L_\beta^T$, $\alpha \geq \beta$ is assumed. Also $\gamma_1 \neq \gamma_2$ is assumed.

KCF : $M_i \oplus M_j$	Z_{ij}^A	Z_{ij}^B	Z_{ji}^A	Z_{ji}^B
$L_\alpha \oplus L_\beta$	0	0	$S^B_{\beta\times(\alpha+1)}$	$-C_\beta^T S^B_{\beta\times(\alpha+1)}$
$L_\alpha \oplus J_\beta(0)$	0	0	$S^L_{\beta\times(\alpha+1)}$	$-C_\beta^T S^L_{\beta\times(\alpha+1)}$
$L_\alpha \oplus J_\beta(\gamma)$	0	0	$F^D_{\beta\times(\alpha+2)} G^T_{\alpha+1}$	$-F^D_{\beta\times(\alpha+2)} \hat{G}^T_{\alpha+1}$
$L_\alpha \oplus N_\beta$	0	0	$C_\beta^T H^L_{\beta\times(\alpha+1)}$	$-H^L_{\beta\times(\alpha+1)}$
$L_\alpha \oplus L_\beta^T$	0	0	$G_{\beta+1} H_{(\beta+2)\times(\alpha+1)}$	$\hat{G}_{\beta+1} H_{(\beta+2)\times(\alpha+1)}$
$J_\alpha(0) \oplus J_\beta(0)$	$S^L_{\alpha\times\beta}$	$-C_\alpha^T S^L_{\alpha\times\beta}$	$T^L_{\beta\times\alpha}$	$-C_\beta^T T^L_{\beta\times\alpha}$
$J_\alpha(0) \oplus L_\beta^T$	0	0	$H^U_{(\beta+1)\times\alpha}$	$-H^U_{(\beta+1)\times\alpha} C_\alpha^T$
$J_\alpha(\gamma) \oplus J_\beta(\gamma)$	$D^L_{\alpha\times\beta}$	$E^L_{\alpha\times\beta}$	$D^L_{\beta\times\alpha}$	$E^L_{\beta\times\alpha}$
$J_\alpha(\gamma) \oplus L_\beta^T$	0	0	$G_{\beta+1}(\hat{I}_\alpha F^D_{\alpha\times(\beta+2)})^T$	$-\hat{G}_{\beta+1}(\hat{I}_\alpha F^D_{\alpha\times(\beta+2)})^T$
$N_\alpha \oplus N_\beta$	$C_\alpha^T S^L_{\alpha\times\beta}$	$-S^L_{\alpha\times\beta}$	$C_\beta^T T^L_{\beta\times\alpha}$	$-T^L_{\beta\times\alpha}$
$N_\alpha \oplus L_\beta^T$	0	0	$T^L_{(\beta+1)\times\alpha} C_\alpha^T$	$-T^L_{(\beta+1)\times\alpha}$
$L_\alpha^T \oplus L_\beta^T$	0	0	$T^B_{(\beta+1)\times\alpha}$	$-T^B_{(\beta+1)\times\alpha} C_\alpha$
$J_\alpha(0) \oplus J_\beta(\gamma)$	0	0	0	0
$J_\alpha(0) \oplus N_\beta$	0	0	0	0
$J_\alpha(\gamma_1) \oplus J_\beta(\gamma_2)$	0	0	0	0
$J_\alpha(\gamma) \oplus N_\beta$	0	0	0	0

Table 2: The diagonal blocks in $Z_A - \lambda Z_B \in \operatorname{nor}(A - \lambda B)$.

KCF : M_i	Z_{ii}^A	Z_{ii}^B
L_α	0	0
$J_\alpha(0)$	$S^L_{\alpha\times\alpha}$	$-C_\alpha^T S^L_{\alpha\times\alpha}$
$J_\alpha(\gamma)$	$D^L_{\alpha\times\alpha}$	$E^L_{\alpha\times\alpha}$
N_α	$C_\alpha^T S^L_{\alpha\times\alpha}$	$-S^L_{\alpha\times\alpha}$
L_α^T	0	0

The superscripts B, L, U, and D of the matrices in tables 1 and 2 are parts of the matrix definitions in Section 5.2. The superscript T is matrix transpose. All subscripts, e.g. $\alpha\times\beta$, refer to the sizes of the matrices.

Notice that the diagonal blocks (i,i) and (j,j) of $Z_A-\lambda Z_B$ can also be obtained from Table 1 by setting $i=j$. For clarity we also display the expressions for the (i,i) and (j,j) blocks of $Z_A-\lambda Z_B$ corresponding to all kinds of structure blocks M_i in Table 2. Of course, the (j,j) blocks corresponding to M_j are read from Table 2 by substituting α with β.

The proof of Theorem 5.1 consists of three parts:

1. The blocks of $Z_A-\lambda Z_B$ displayed in Table 1 fulfill the conditions in Proposition 5.1, which imply that $Z_A-\lambda Z_B\in\mathrm{nor}(A-\lambda B)$ is orthogonal to an arbitrary $T_A-\lambda T_B\in\tan(A-\lambda B)$.

2. The number of independent parameters in $Z_A-\lambda Z_B$ is equal to the codimension of $\mathrm{orbit}(A-\lambda B)$, which implies that the parameterized normal form has minimum number of parameters.

3. Each block in Table 1 defines an orthogonal basis, i.e., the basis for each parameter p_i is orthogonal to the basis for each other parameter p_j, $i\neq j$.

We start by proving part 3, followed by proving parts 1 and 2 for the 16 different cases $\mathrm{diag}(M_i,M_j)$ corresponding to different combinations of structure blocks in the KCF. In Table 3 we display the codimension for these 16 cases and the number of parameters in the $(i,i),(i,j),(j,i)$ and (j,j) blocks of $Z_A-\lambda Z_B$. The codimensions are computed from (4.10), which is the minimum number of parameters required to span the corresponding normal space. For the ordering and the sizes of the blocks in $A-\lambda B$ we have made the same assumptions in Table 3 as in Table 1. Notice that the codimension counts for $L_\alpha\oplus L_\beta$ and $L_\alpha^T\oplus L_\beta^T$ are 0 if $\alpha=\beta$. The number of parameters required in each of the $(i,i),(i,j),(j,i)$ and (j,j) blocks of $Z_A-\lambda Z_B$ follows from the proof given below.

Proof of part 3. We show that each matrix pencil block in Table 1 has all its parameters in orthogonal directions. This is trivial for blocks built from the structured Toeplitz and Hankel matrices S^L, S^B, H, H^L, H^U, T^L, or T^B (possibly involving some kind of shift). Remember that the Frobenius inner product can be expressed in terms of the sum of all results from elementwise multiplications as shown in (5.6). For each of these matrices, the elementwise multiplication of the basis for one parameter p_i and the basis for another parameter $p_j, j\neq i$ only results in multiplications where at least one of the two elements is zero. Obviously, these bases are orthogonal. For the matrix pencil blocks built from the F^D matrix, the orthogonality follows from construction, since some of the elements are explicitly chosen so that the Frobenius inner product is zero.

For the proof for the blocks of type $D^L-\lambda E^L$ we define s_q in terms of the d_i and e_i in (5.8) to be

$$s_q=\sum_{i=1}^{q} i|d_i|^2+\sum_{i=1}^{q-1} i|e_i|^2-q\overline{\gamma}d_q\overline{e}_q.$$

Independent of s and t, the number s_q is the inner product of the qth basis vector with the rth, where $q<r$.

We show by induction that $s_q=0$ for $q=1,2,\ldots$. Clearly $s_1=|\gamma|^2-\gamma\overline{\gamma}=0$.

Table 3: The number of parameters in the (i,i), (i,j), (j,i), and (j,j) blocks of $Z_A - \lambda Z_B \in \mathrm{nor}(M_i \oplus M_j)$.

KCF:$M_i \oplus M_j$	$\mathrm{cod}(M_i \oplus M_j)$	(i,i)	(i,j)	(j,i)	(j,j)
$L_\alpha \oplus L_\beta$	$\beta-\alpha-1$	0	0	$\beta-\alpha-1$	0
$L_\alpha \oplus J_\beta(0)$	2β	0	0	β	β
$L_\alpha \oplus J_\beta(\gamma)$	2β	0	0	β	β
$L_\alpha \oplus N_\beta$	2β	0	0	β	β
$L_\alpha \oplus L_\beta^T$	$\alpha+\beta+2$	0	0	$\alpha+\beta+2$	0
$J_\alpha(0) \oplus J_\beta(0)$	$\beta+3\alpha$	α	α	α	β
$J_\alpha(0) \oplus L_\beta^T$	2α	α	0	α	0
$J_\alpha(\gamma) \oplus J_\beta(\gamma)$	$\beta+3\alpha$	α	α	α	β
$J_\alpha(\gamma) \oplus L_\beta^T$	2α	α	0	α	0
$N_\alpha \oplus N_\beta$	$\beta+3\alpha$	α	α	α	β
$N_\alpha \oplus L_\beta^T$	2α	α	0	α	0
$L_\alpha^T \oplus L_\beta^T$	$\alpha-\beta-1$	0	0	$\alpha-\beta-1$	0
$J_\alpha(0) \oplus J_\beta(\gamma)$	$\alpha+\beta$	α	0	0	β
$J_\alpha(0) \oplus N_\beta$	$\alpha+\beta$	α	0	0	β
$J_\alpha(\gamma_1) \oplus J_\beta(\gamma_2)$	$\alpha+\beta$	α	0	0	β
$J_\alpha(\gamma) \oplus N_\beta$	$\alpha+\beta$	α	0	0	β

We now show that $s_{q+1} - s_q = 0$ from which the result follows.

$$
\begin{aligned}
&q\overline{\gamma}d_q\overline{e}_q + (q+1)|d_{q+1}|^2 + q|e_q|^2 - (q+1)\overline{\gamma}d_{q+1}\overline{e}_{q+1} = \\
&q\overline{e}_q(\overline{\gamma}d_q + e_q) + (q+1)d_{q+1}(\overline{d}_{q+1} - \overline{\gamma}\ \overline{e}_{q+1}) = \\
&d_{q+1}((q+1)(\overline{d}_{q+1} - \overline{\gamma}\ \overline{e}_{q+1}) - q\overline{e}_q) = \\
&d_{q+1}((q+1)(-\gamma\overline{d}_q - \overline{e}_q + \gamma\overline{d}_q + 2\overline{e}_q - \frac{\overline{e}_q}{q+1}) - q\overline{e}_q) = \\
&d_{q+1}((q+1)\overline{e}_q - \overline{e}_q - q\overline{e}_q) = 0.
\end{aligned}
$$

Since $Z_A - \lambda Z_B$ is built from b^2 mutually independent blocks in Table 1, each associated with c_i parameters, it follows that $Z_A - \lambda Z_B$ is an orthogonal basis for a $c_1 + c_2 + \ldots + c_{b^2}$ dimensional space, with one parameter for each dimension. □

Proof of parts 1 and 2. Now, it remains to show that $Z_A - \lambda Z_B$ is orthogonal to $\tan(A - \lambda B)$ and that the number of parameters in $Z_A - \lambda Z_B$ is equal to $\mathrm{cod}(A - \lambda B)$. Since the number of parameters in orthogonal directions cannot exceed the codimension, it is sufficient to show that we have found them all. The orthogonality between $Z_A - \lambda Z_B$ and $\tan(A - \lambda B)$ is shown by proving that each pair of blocks fulfills the conditions $A_j^H Z_{ji}^A = -B_j^H Z_{ji}^B$ and $Z_{ji}^A A_i^H = -Z_{ji}^B B_i^H$ in Proposition 5.1. In the following we refer to these as the *first* and *second* conditions, respectively.

We carry out the proofs for all 16 cases $M_i \oplus M_j$ in Table 1, starting with blocks where M_i and M_j are of the same kind.

$\mathbf{J}_\alpha(\mathbf{0}) \oplus \mathbf{J}_\beta(\mathbf{0})$: We note that $J_k(0) = C_k - \lambda I_k$. *First* condition for the (j,i) block:

$$A_j^H Z_{ji}^A = C_\beta^T T_{\beta\times\alpha}^L = I_\beta C_\beta^T T_{\beta\times\alpha}^L = -B_j^H Z_{ji}^B.$$

Second condition for the (j,i) block:

$$Z_{ji}^A A_i^H = T_{\beta\times\alpha}^L C_\alpha^T = T_{\beta\times\alpha}^L C_\alpha^T I_\alpha = C_\beta^T T_{\beta\times\alpha}^L I_\alpha = -Z_{ji}^B B_i^H,$$

where we used that $T_{\beta\times\alpha}^L C_\alpha^T = C_\beta^T T_{\beta\times\alpha}^L$, for $\beta \geq \alpha$. Similarly for the (i,j) block:

$$A_i^H Z_{ij}^A = C_\alpha^T S_{\alpha\times\beta}^L = I_\alpha C_\alpha^T S_{\alpha\times\beta}^L = -B_i^H Z_{ij}^B,$$

and

$$Z_{ij}^A A_j^H = S_{\alpha\times\beta}^L C_\beta^T = S_{\alpha\times\beta}^L C_\beta^T I_\beta = C_\alpha^T S_{\alpha\times\beta}^L I_\beta = -Z_{ij}^B B_j^H.$$

Here we used that $S_{\alpha\times\beta}^L C_\beta^T = C_\alpha^T S_{\alpha\times\beta}^L$, for $\beta \geq \alpha$.

Since the (i,i), (i,j), and (j,i) blocks of $Z_A - \lambda Z_B$, have α parameters each and the (j,j) block has β parameters, the total number of parameters in $Z_A - \lambda Z_B$ is equal to $\mathrm{cod}(J_\alpha(0) \oplus J_\beta(0)) = \beta + 3\alpha$.

$\mathbf{N}_\alpha \oplus \mathbf{N}_\beta$: Since there is a symmetry between $J_k(0) = C_k - \lambda I_k$ and $N_k = I_k - \lambda C_k$ and there is a corresponding symmetry between blocks in $Z_A - \lambda Z_B$ for $J_k(0)$ and N_k blocks, the proof for $N_\alpha \oplus N_\beta$ is similar to the case $J_\alpha(0) \oplus J_\beta(0)$.

$\mathbf{J}_\alpha(\gamma) \oplus \mathbf{J}_\beta(\gamma)$: Here the (j,i) block and the (i,j) block are defined similarly (see Table 1), and therefore it is sufficient to prove one of them with no constraints on α and β. We note that $J_k(\gamma) = \gamma I_k + C_k - \lambda I_k$. We show that the first and second conditions hold for $Z_{ji}^A = D_{\beta\times\alpha}[q]$ and $Z_{ji}^B = E_{\beta\times\alpha}[q]$ for $q = 1, \ldots, \min\{\alpha, \beta\}$. *First* condition:

$$A_j^H Z_{ji}^A = (\gamma I_\beta + C_\beta)^H D_{\beta\times\alpha}[q] = \overline{\gamma} D_{\beta\times\alpha}[q] + C_\beta^T D_{\beta\times\alpha}[q].$$

Remember that $D_{\beta\times\alpha}[q]$ has all elements zero, except for the q lower left diagonals, where all elements in each diagonal are identical and defined by the element in the first column. For $q = 1$ the proof is trivial. For $q > 1$, $A_j^H Z_{ji}^A$ gives the following matrix. All diagonals starting at position $(u, 1)$ for $1 \leq u \leq \beta - q$ are zero. The elements in the diagonal starting at position $(\beta - q + 1, 1)$ are $\overline{\gamma} d_q$ which by definition is equal to $-e_q^*$, which in turn defines the corresponding diagonal in $-E_{\beta\times\alpha}[q]$. The elements in the diagonals starting at positions $(\beta - u + 1, 1)$, where $1 \leq u < q$ are equal to $\overline{\gamma} d_u + d_{u+1}$. Since d_{u+1} is defined as $-\overline{\gamma} d_u - e_u$, the elements in these diagonals are equal to $-e_u$, which defines the elements in the corresponding diagonals in $-E_{\beta\times\alpha}[q]$. Since $-E_{\beta\times\alpha}[q] = -B_j^H Z_{ji}^B$, we have proved the first condition.

Second condition: Since $D_{\beta\times\alpha}[q]$ only has $q \leq \min\{s, t\}$ non-zero diagonals in the lower left corner of the matrix, a shift of rows downwards gives the same result as a shift of columns leftwards, i.e., $C_\beta^T D_{\beta\times\alpha}[q] = D_{\beta\times\alpha}[q] C_\alpha^T$. Using information from the first part we obtain

$$\begin{aligned} Z_{ji}^A A_i^H &= D_{\beta\times\alpha}[q](\gamma I_\alpha + C_\alpha)^H = \overline{\gamma} D_{\beta\times\alpha}[q] + D_{\beta\times\alpha}[q] C_\alpha^T = \overline{\gamma} D_{\beta\times\alpha}[q] + C_\beta^T D_{\beta\times\alpha}[q] \\ &= A_j^H Z_{ji}^A = -E_{\beta\times\alpha}[q] = -Z_{ji}^B B_i^H, \end{aligned}$$

since B_i is the identity matrix.

Also here, the number of parameters in $Z_{jj}^A - \lambda Z_{jj}^B$ is β and it is α parameters in each of the other three blocks, giving $\beta + 3\alpha$ in total.

Even though the (i,i), (j,i), (i,j), and (j,j) blocks look rather complicated, they reduce for $\gamma = 0$ to the corresponding blocks for $J_\alpha(0) \oplus J_\beta(0)$ in Table 1.

$\mathbf{L}_\alpha \oplus \mathbf{L}_\beta$: Here we use $L_k = \hat{G}_k - \lambda G_k$. *First* condition for the (j,i) block:

$$A_j^H Z_{ji}^A = \hat{G}_\beta^T S_{\beta\times(\alpha+1)}^B = \begin{bmatrix} 0 \\ S_{\beta\times(\alpha+1)}^B \end{bmatrix} = \begin{bmatrix} C_\beta^T S_{\beta\times(\alpha+1)}^B \\ 0 \end{bmatrix} = G_\beta^T C_\beta^T S_{\beta\times(\alpha+1)}^B = -B_j^H Z_{ji}^B.$$

Second condition for the (j,i) block:

$$Z_{ji}^A A_i^H = S_{\beta\times(\alpha+1)}^B \hat{G}_\beta^T = \begin{bmatrix} 0 \\ S_{\beta\times(\alpha+1)}^B \end{bmatrix} = \begin{bmatrix} C_\beta^T S_{\beta\times(\alpha+1)}^B \\ 0 \end{bmatrix} = C_\beta^T S_{\beta\times(\alpha+1)}^B G_\beta^T = -Z_{ji}^B B_i^H.$$

Since the contribution from $L_\alpha \oplus L_\beta$ to the codimension is $\beta - \alpha - 1$ and the (j,i) block has $\beta - \alpha - 1$ independent parameters we deduce that all other blocks in $Z_A - \lambda Z_B$ are zero.

$\mathbf{L}_\alpha^\mathbf{T} \oplus \mathbf{L}_\beta^\mathbf{T}$: Since this case is just the transpose of $L_\alpha \oplus L_\beta$ the proof is almost the same, and therefore we omit the technical details here.

So far, we have proved all cases where both blocks are of the same type. Since the diagonal blocks in $Z_A - \lambda Z_B$ always correspond to such cases (see Table 3 for the number of parameters in these blocks), we from now on only have to consider the (i,j) and (j,i) blocks, where $i \neq j$ for the remaining cases.

$\mathbf{L}_\alpha \oplus \mathbf{J}_\beta(\mathbf{0})$: *First* condition for the (j,i) block:

$$A_j^H Z_{ji}^A = C_\beta^T S_{\beta\times(\alpha+1)}^L = I_\beta C_\beta^T S_{\beta\times(\alpha+1)}^L = -B_j^H Z_{ji}^B.$$

Second condition for the (j,i) block:

$$Z_{ji}^A A_i^H = S_{\beta\times(\alpha+1)}^L \hat{G}_\alpha^T = C_\beta^T S_{\beta\times(\alpha+1)}^L G_\alpha^T = -Z_{ji}^B B_i^H.$$

The (i,i) and (j,j) blocks contribute with zero and β parameters, respectively. Since the (j,i) block gives another β parameters, we have found all 2β parameters, and therefore it follows that $Z_{ij}^A = \lambda Z_{ij}^B = 0$.

$\mathbf{L}_\alpha \oplus \mathbf{J}_\beta(\gamma)$: *First* condition for the (j,i) block:

$$A_j^H Z_{ji}^A = (\gamma I_\beta + C_\beta)^H F_{\beta\times(\beta+2)}^D G_{\alpha+1}^T = \overline{\gamma} F_{\beta\times(\beta+2)}^D G_{\alpha+1}^T + C_\beta^T F_{\beta\times(\beta+2)}^D G_{\alpha+1}^T.$$

By inspection we see that the (u,v)-element of this matrix is $\overline{\gamma} f_{u,v}^d + f_{u-1,v}^d$ if $u > 1$ and $\overline{\gamma} f_{u,v}^d$ if $u = 1$ (where $f_{u,v}^d$ denotes the (u,v)-element of F^D). The right hand side of the same condition is

$$-B_j^H Z_{ji}^B = I_\beta F_{\beta\times(\beta+2)}^D \hat{G}_{\alpha+1}^T,$$

which simply is the β leftmost columns of $F_{\beta\times(\beta+2)}^D$. The (u,v)-element of this matrix is then $f_{u,v+1}^d$, which is defined as $\overline{\gamma} f_{u,v}^d + f_{u-1,v}^d$ if $u > 1$ and $\overline{\gamma} f_{u,v}^d$ if $u = 1$.

Second condition for the (j,i) block:

$$Z_{ji}^A A_i^H \hat{G}_\alpha^T = F_{\beta\times(\alpha+2)}^D G_{\alpha+1}^T \hat{G}_\alpha^T = F_{\beta\times(\alpha+2)}^D \begin{bmatrix} 0 \\ I_\alpha \\ 0 \end{bmatrix} = F_{\beta\times(\alpha+2)}^D \hat{G}_{\alpha+1} G_\alpha^T = -Z_{ji}^B B_i^H.$$

As in the previous case, the (i,i) and (j,j) blocks contribute with zero and β parameters, respectively. Since the (j,i) block gives the remaining β parameters, the (i,j) block is the zero pencil.

Notably, for $\gamma = 0$, the "monstrous" (j,i) block reduces to the (j,i) block for $L_\alpha \oplus J_\beta(0)$ in Table 1.

$\mathbf{L}_\alpha \oplus \mathbf{N}_\beta$: *First* condition for the (j,i) block:

$$A_j^H Z_{ji}^A = I_\beta C_\beta^T H_{\beta\times(\alpha+1)}^L = C_\beta^T H_{\beta\times(\alpha+1)}^L = -B_j^H Z_{ji}^B.$$

Second condition for the (j,i) block:

$$Z_{ji}^A A_i^H \hat{G}_\alpha^T = C_\beta^T H_{\beta\times(\alpha+1)}^L = \begin{bmatrix} 0 \\ H_{(\beta-1)\times\alpha}^L \end{bmatrix} = H_{\beta\times(\alpha+1)}^L G_\alpha^T = -Z_{ji}^B B_i^H.$$

Also here, the (i,i) and (j,j) blocks contribute with zero and β parameters, respectively. Since the (j,i) block gives the remaining β parameters, the (i,j) block is the zero pencil.

$\mathbf{L}_\alpha \oplus \mathbf{L}_\beta^{\mathbf{T}}$: For this case the (i,i) and (j,j) blocks are zero pencils. *First* condition for the (j,i) block:

$$A_j^H Z_{ji}^A = \hat{G}_\beta G_{\beta+1} H_{(\beta+2)\times(\alpha+1)} = [0 \ I_\beta \ 0] H_{(\beta+2)\times(\alpha+1)} = G_\beta \hat{G}_{\beta+1} H_{(\beta+2)\times(\alpha+1)} = -B_j^H Z_{ji}^B.$$

Second condition for the (j,i) block:

$$Z_{ji}^A A_i^H = G_{\beta+1} H_{(\beta+2)\times(\alpha+1)} \hat{G}_\alpha^T,$$

which is a matrix consisting of the $\beta+1$ first rows and α last columns of $H_{(\beta+2)\times(\alpha+1)}$. This matrix is identical to the one given by the $\beta+1$ last rows and α first columns of $H_{(\beta+2)\times(\alpha+1)}$, i.e.,

$$\hat{G}_{\beta+1} H_{(\beta+2)\times(\alpha+1)} G_\alpha^T = -Z_{ji}^B B_i^H.$$

Since this block has all $\alpha+\beta+2$ parameters, it follows that the (i,j) block is the zero pencil.

$\mathbf{J}_\alpha(\mathbf{0}) \oplus \mathbf{L}_\beta^{\mathbf{T}}$: *First* condition for the (j,i) block:

$$A_j^H Z_{ji}^A = \hat{G}_\beta H_{(\beta+1)\times\alpha}^U,$$

which simply is the last β rows in $H_{(\beta+1)\times\alpha}^U$. Another way to construct this matrix is to shift the columns in $H_{(\beta+1)\times\alpha}^U$ one column leftwards and pick the β first columns of that matrix, which can be written as

$$G_\beta H_{(\beta+1)\times\alpha}^U C_\alpha^T = -B_j^H Z_{ji}^B.$$

Second condition for the (j, i) block:

$$Z_{ji}^A A_i^H = H_{(\beta+1)\times\alpha}^U C_\alpha^T = H_{(\beta+1)\times\alpha}^U C_\alpha^T I_\alpha = -Z_{ji}^B B_i^H.$$

The (i, i) and (j, j) blocks contribute with α and zero parameters, respectively. Since the (j, i) block gives another α parameters, we conclude that the (i, j) block is the zero pencil.

$\mathbf{J}_\alpha(\gamma) \oplus \mathbf{L}_\beta^\mathbf{T}$: Since the proof for this case is similar to the one for the case $L_\alpha \oplus J_\beta(\gamma)$, we omit the technical details here. It follows that for $\gamma = 0$, the (j, i) block reduces to the (j, i) block for $J_\alpha(0) \oplus L_\beta^T$ in Table 1.

$\mathbf{N}_\alpha \oplus \mathbf{L}_\beta^\mathbf{T}$: *First* condition for the (j, i) block:

$$A_j^H Z_{ji}^A = \hat{G}_\beta T_{(\beta+1)\times\alpha}^L C_\alpha^T,$$

which is the last β rows in $T_{(\beta+1)\times\alpha}^L$ shifted one column leftwards. This matrix is identical to the one given by the β first rows in $T_{(\beta+1)\times\alpha}^L$, which is

$$G_\beta T_{(\beta+1)\times\alpha}^L = -B_j^H Z_{ji}^B.$$

Second condition for the (j, i) block:

$$Z_{ji}^A A_i^H = T_{(\beta+1)\times\alpha}^L C_\alpha^T I_\alpha = T_{(\beta+1)\times\alpha}^L C_\alpha^T = -Z_{ji}^B B_i^H.$$

The (i, i) and (j, j) blocks in $Z_A - \lambda Z_B$ contribute with α and zero parameters, respectively. Since the (j, i) block gives another α parameters, we conclude that the (i, j) block is the zero pencil.

$\mathbf{J}_\alpha(\mathbf{0}) \oplus \mathbf{J}_\beta(\gamma)$, $\mathbf{J}_\alpha(\mathbf{0}) \oplus \mathbf{N}_\beta$, $\mathbf{J}_\alpha(\gamma_1) \oplus \mathbf{J}_\beta(\gamma_2)$, and $\mathbf{J}_\alpha(\gamma) \oplus \mathbf{N}_\beta$: In these four cases the (i, i) and (j, j) blocks contribute with α and β parameters, respectively, and therefore the (j, i) and (i, j) blocks are zero pencils.

Since we have considered all possible cases of M_i and M_j blocks the proof is complete. □

6 Applications and Examples

6.1 Some examples of versal deformations of matrix pencils in KCF

In the following we show three examples of versal deformations of matrix pencils. For the 7×8 pencil $A - \lambda B = L_2 \oplus J_2(0) \oplus J_3(0)$ with codimension 14, the 14-parameter versal deformation $A - \lambda B + Z_A - \lambda Z_B$, where $Z_A - \lambda Z_B \in \mathrm{nor}(A - \lambda B)$ is given by

$$Z_A = \left[\begin{array}{ccc|cc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline p_1 & 0 & 0 & p_6 & 0 & p_{10} & 0 & 0 \\ p_2 & p_1 & 0 & p_7 & p_6 & p_{11} & p_{10} & 0 \\ \hline p_3 & 0 & 0 & 0 & 0 & p_{12} & 0 & 0 \\ p_4 & p_3 & 0 & p_8 & 0 & p_{13} & p_{12} & 0 \\ p_5 & p_4 & p_3 & p_9 & p_8 & p_{14} & p_{13} & p_{12} \end{array}\right],$$

and

$$
Z_B = \left[\begin{array}{ccc|cc|ccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-p_1 & 0 & 0 & -p_6 & 0 & -p_{10} & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-p_3 & 0 & 0 & 0 & 0 & -p_{12} & 0 & 0 \\
-p_4 & -p_3 & 0 & -p_8 & 0 & -p_{13} & -p_{12} & 0
\end{array}\right].
$$

For the 3×4 pencil $A - \lambda B = L_1 \oplus J_2(\gamma)$ with codimension 4, the 4-parameter versal deformation $A - \lambda B + Z_A - \lambda Z_B$, where $Z_A - \lambda Z_B \in \operatorname{nor}(A - \lambda B)$ is given by

$$
Z_A = \left[\begin{array}{cc|cc}
0 & 0 & 0 & 0 \\
\hline
p_1 & p_1\overline{\gamma} & p_3(|\gamma|^2+1) & 0 \\
p_2 - p_1\frac{2\gamma}{|\gamma|^2+1} & p_2\overline{\gamma} - p_1\frac{|\gamma|^2-1}{|\gamma|^2+1} & -p_3\gamma + p_4 & p_3(|\gamma|^2+1)
\end{array}\right],
$$

and

$$
Z_B = \left[\begin{array}{cc|cc}
0 & 0 & 0 & 0 \\
\hline
-p_1\overline{\gamma} & -p_1\overline{\gamma}^2 & -p_3(|\gamma|^2\overline{\gamma}+\overline{\gamma}) & 0 \\
-p_2\overline{\gamma} + p_1\frac{|\gamma|^2-1}{|\gamma|^2+1} & -p_2\overline{\gamma}^2 - p_1\frac{2\overline{\gamma}}{|\gamma|^2+1} & -p_3 - p_4\overline{\gamma} & -p_3(|\gamma|^2\overline{\gamma}+\overline{\gamma})
\end{array}\right].
$$

For the 11×11 pencil $A - \lambda B = L_1 \oplus J_3(0) \oplus N_4 \oplus L_2^T$ with codimension 26, the 26-parameter versal deformation $A - \lambda B + Z_A - \lambda Z_B$, where $Z_A - \lambda Z_B \in \operatorname{nor}(A - \lambda B)$ is given by

$$
Z_A = \left[\begin{array}{cc|ccc|cccc|cc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
p_1 & 0 & p_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
p_2 & p_1 & p_{14} & p_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
p_3 & p_2 & p_{15} & p_{14} & p_{13} & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & p_4 & 0 & 0 & 0 & p_{19} & 0 & 0 & 0 & 0 & 0 \\
p_4 & p_5 & 0 & 0 & 0 & p_{20} & p_{19} & 0 & 0 & 0 & 0 \\
p_5 & p_6 & 0 & 0 & 0 & p_{21} & p_{20} & p_{19} & 0 & 0 & 0 \\
\hline
p_8 & p_9 & p_{18} & p_{17} & p_{16} & p_{23} & 0 & 0 & 0 & 0 & 0 \\
p_9 & p_{10} & p_{17} & p_{16} & 0 & p_{24} & p_{23} & 0 & 0 & 0 & 0 \\
p_{10} & p_{11} & p_{16} & 0 & 0 & p_{25} & p_{24} & p_{23} & 0 & 0 & 0
\end{array}\right],
$$

and

$$Z_B = \left[\begin{array}{cc|ccc|cccc|cc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -p_1 & 0 & -p_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -p_2 & -p_1 & -p_{14} & -p_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & -p_4 & 0 & 0 & 0 & -p_{19} & 0 & 0 & 0 & 0 & 0 \\ -p_4 & -p_5 & 0 & 0 & 0 & -p_{20} & -p_{19} & 0 & 0 & 0 & 0 \\ -p_5 & -p_6 & 0 & 0 & 0 & -p_{21} & -p_{20} & -p_{19} & 0 & 0 & 0 \\ -p_6 & -p_7 & 0 & 0 & 0 & -p_{22} & -p_{21} & -p_{20} & -p_{19} & 0 & 0 \\ \hline -p_9 & -p_{10} & -p_{17} & -p_{16} & 0 & -p_{24} & -p_{23} & 0 & 0 & 0 & 0 \\ -p_{10} & -p_{11} & -p_{16} & 0 & 0 & -p_{25} & -p_{24} & -p_{23} & 0 & 0 & 0 \\ -p_{11} & -p_{12} & 0 & 0 & 0 & -p_{26} & -p_{25} & -p_{24} & -p_{23} & 0 & 0 \end{array}\right].$$

6.2 Versal deformations of the set of 2-by-3 matrix pencils

In [14], the algebraic and geometric characteristics of the set of 2-by-3 matrix pencils were examined in full detail, including the complete closure hierarchy. There, all non-zero and finite eigenvalues were considered as unspecified. R_2 was used to denote a 2-by-2 block with non-zero finite eigenvalues, i.e., any of the three structures $J_1(\alpha)\oplus J_1(\beta)$, $J_1(\alpha)\oplus J_1(\alpha)$, and $J_2(\alpha)$, where $\alpha, \beta \neq \{0, \infty\}$. However, in the context of versal deformations all these forms are considered separately and with the eigenvalues specified (known). Consequently, we now have 20 different Kronecker structures to investigate. For example, the versal deformation of $A - \lambda B = L_0 \oplus J_2(\gamma)$, $\gamma \neq \{0, \infty\}$, is found by computing $Z_A - \lambda Z_B =$

$$\begin{bmatrix} p_1 + \lambda\,\overline{\gamma}\,p_1 & p_3(|\gamma|^2+1) + p_3(|\gamma|^2\overline{\gamma}+\overline{\gamma}) & 0 \\ p_2 - \frac{p_1\gamma}{|\gamma|^2+1} + \lambda(p_2\overline{\gamma} + \frac{p_1}{|\gamma|^2+1}) & -p_3\gamma + p_4 + \lambda(p_3 + p_4\overline{\gamma}) & p_3(|\gamma|^2+1) + p_3(|\gamma|^2\overline{\gamma}+\overline{\gamma}) \end{bmatrix}. \quad (6.1)$$

In Table 4 we show the versal deformations for all different Kronecker structures for this set of matrix pencils. The different structures are displayed in increasing codimension order.

6.2.1 Using GUPTRI in a random walk in tangent and normal directions of non-generic pencils

In order to illustrate how perturbations in the tangent space and in the normal space affect the Kronecker structure computed by a staircase algorithm, we have performed a set of tests on non-generic 2-by-3 matrix pencils. Since the staircase algorithm considers all non-zero finite eigenvalues as unspecified, we have not included these cases in the test.

For the remaining 12 non-generic cases a random perturbation $E_A - \lambda E_B$, with entries e^a_{ij}, e^b_{ij}, has been decomposed into two parts, $T_A - \lambda T_B \in \tan(A - \lambda B)$, and $Z_A - \lambda Z_B \in \operatorname{nor}(A - \lambda B)$, such that

$$E_A = T_A + Z_A \quad \text{and} \quad E_B = T_B + Z_B.$$

We illustrate the decomposition of $E_A - \lambda E_B$ with $A - \lambda B = L_0 \oplus J_2(0)$. From Table 4 we get

$$Z_A = \begin{bmatrix} p_1 & p_3 & 0 \\ p_2 & p_4 & p_3 \end{bmatrix}, \quad Z_B = \begin{bmatrix} 0 & 0 & 0 \\ -p_1 & -p_3 & 0 \end{bmatrix}.$$

Table 4: Versal deformations $A - \lambda B + Z_A - \lambda Z_B$ of the set of 2-by-3 matrix pencils.

KCF	$A - \lambda B$	$Z_A - \lambda Z_B$
L_2	$\begin{bmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
$L_1 \oplus J_1(\gamma)$	$\begin{bmatrix} -\lambda & 1 & 0 \\ 0 & 0 & \gamma - \lambda \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 \\ p_1 + \lambda \overline{\gamma}\, p_1 & \overline{\gamma}\, p_1 + \lambda \overline{\gamma}^2 p_1 & p_2 + \lambda \overline{\gamma}\, p_2 \end{bmatrix}$
$L_1 \oplus J_1(0)$	$\begin{bmatrix} -\lambda & 1 & 0 \\ 0 & 0 & -\lambda \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 \\ p_1 & 0 & p_2 \end{bmatrix}$
$L_1 \oplus N_1$	$\begin{bmatrix} -\lambda & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda\, p_1 & \lambda\, p_2 \end{bmatrix}$
$L_0 \oplus J_1(\gamma_1) \oplus J_1(\gamma_2)$	$\begin{bmatrix} 0 & \gamma_1 - \lambda & 0 \\ 0 & 0 & \gamma_2 - \lambda \end{bmatrix}$	$\begin{bmatrix} p_1 + \lambda \overline{\gamma}_1\, p_1 & p_3 + \lambda \overline{\gamma}_1\, p_3 & 0 \\ p_2 + \lambda \overline{\gamma}_2\, p_2 & 0 & p_4 + \lambda \overline{\gamma}_2\, p_4 \end{bmatrix}$
$L_0 \oplus J_2(\gamma)$	$\begin{bmatrix} 0 & \gamma - \lambda & 1 \\ 0 & 0 & \gamma - \lambda \end{bmatrix}$	See (6.1)
$L_0 \oplus 2J_1(\gamma)$	$\begin{bmatrix} 0 & \gamma - \lambda & 0 \\ 0 & 0 & \gamma - \lambda \end{bmatrix}$	$\begin{bmatrix} p_1 + \lambda \overline{\gamma}\, p_1 & p_3 + \lambda \overline{\gamma}\, p_3 & p_5 + \lambda \overline{\gamma}\, p_5 \\ p_2 + \lambda \overline{\gamma}\, p_2 & p_4 + \lambda \overline{\gamma}\, p_4 & p_6 + \lambda \overline{\gamma}\, p_6 \end{bmatrix}$
$L_0 \oplus J_1(0) \oplus J_1(\gamma)$	$\begin{bmatrix} 0 & -\lambda & 0 \\ 0 & 0 & \gamma - \lambda \end{bmatrix}$	$\begin{bmatrix} p_1 & p_3 & 0 \\ p_2 + \lambda \overline{\gamma}\, p_2 & 0 & p_4 + \lambda \overline{\gamma}\, p_4 \end{bmatrix}$
$L_0 \oplus J_1(\gamma) \oplus N_1$	$\begin{bmatrix} 0 & \gamma - \lambda & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} p_1 + \lambda \overline{\gamma}\, p_1 & p_3 + \lambda \overline{\gamma}\, p_3 & 0 \\ \lambda\, p_2 & 0 & \lambda\, p_4 \end{bmatrix}$
$L_0 \oplus J_2(0)$	$\begin{bmatrix} 0 & -\lambda & 1 \\ 0 & 0 & -\lambda \end{bmatrix}$	$\begin{bmatrix} p_1 & p_3 & 0 \\ p_2 + \lambda\, p_1 & p_4 + \lambda\, p_3 & p_3 \end{bmatrix}$
$L_0 \oplus N_2$	$\begin{bmatrix} 0 & 1 & -\lambda \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} \lambda\, p_1 & \lambda\, p_3 & 0 \\ p_1 + \lambda\, p_2 & p_3 + \lambda\, p_4 & \lambda\, p_3 \end{bmatrix}$
$L_0 \oplus J_1(0) \oplus N_1$	$\begin{bmatrix} 0 & -\lambda & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} p_1 & p_3 & 0 \\ \lambda\, p_2 & 0 & \lambda\, p_4 \end{bmatrix}$
$L_0 \oplus L_1 \oplus L_0^T$	$\begin{bmatrix} 0 & -\lambda & 1 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 \\ p_1 + \lambda\, p_2 & p_3 + \lambda\, p_4 & p_4 + \lambda\, p_5 \end{bmatrix}$
$L_0 \oplus 2J_1(0)$	$\begin{bmatrix} 0 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{bmatrix}$	$\begin{bmatrix} p_1 & p_3 & p_5 \\ p_2 & p_4 & p_6 \end{bmatrix}$
$L_0 \oplus 2N_1$	$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} \lambda\, p_1 & \lambda\, p_3 & \lambda\, p_5 \\ \lambda\, p_2 & \lambda\, p_4 & \lambda\, p_6 \end{bmatrix}$
$2L_0 \oplus L_1^T$	$\begin{bmatrix} 0 & 0 & -\lambda \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} p_1 + \lambda\, p_2 & p_4 + \lambda\, p_5 & 0 \\ p_2 + \lambda\, p_3 & p_5 + \lambda\, p_6 & 0 \end{bmatrix}$
$2L_0 \oplus J_1(\gamma) \oplus L_0^T$	$\begin{bmatrix} 0 & 0 & \gamma - \lambda \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} p_1 + \lambda \overline{\gamma}\, p_1 & p_4 + \lambda \overline{\gamma}\, p_4 & p_7 + \lambda \overline{\gamma}\, p_7 \\ p_2 + \lambda\, p_3 & p_5 + \lambda\, p_6 & p_8 + \lambda \overline{\gamma}\, p_8 \end{bmatrix}$
$2L_0 \oplus J_1(0) \oplus L_0^T$	$\begin{bmatrix} 0 & 0 & -\lambda \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} p_1 & p_4 & p_7 \\ p_2 + \lambda\, p_3 & p_5 + \lambda\, p_6 & p_8 \end{bmatrix}$
$2L_0 \oplus N_1 \oplus L_0^T$	$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} \lambda\, p_1 & \lambda\, p_4 & \lambda\, p_7 \\ p_2 + \lambda\, p_3 & p_5 + \lambda\, p_6 & \lambda\, p_8 \end{bmatrix}$
$3L_0 \oplus 2L_0^T$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} p_1 + \lambda\, p_2 & p_5 + \lambda\, p_6 & p_9 + \lambda\, p_{10} \\ p_3 + \lambda\, p_4 & p_7 + \lambda\, p_8 & p_{11} + \lambda\, p_{12} \end{bmatrix}$

Let $T_A - \lambda T_B = (E_A - \lambda E_B) - (Z_A - \lambda Z_B)$. Now, the parameters p_i are determined by computing the component of $E_A - \lambda E_B$ in each of the four orthogonal (but not orthonormal) directions that span the normal space:

$$
\begin{aligned}
Z_1 &= \frac{1}{2}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \right) \\
Z_2 &= 1\left(\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) \\
Z_3 &= \frac{1}{3}\left(\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \right) \\
Z_4 &= 1\left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right).
\end{aligned}
$$

We conclude that

$$
p_1 = \frac{e_{11}^a - e_{21}^b}{2}, \quad p_2 = e_{21}^a, \quad p_3 = \frac{e_{12}^a + e_{23}^a - e_{22}^b}{3}, \quad p_4 = e_{22}^a.
$$

It is easily verified that $< T_A - \lambda T_B, Z_A - \lambda Z_B > = 0$.

GUPTRI [11, 12] has been used to compute the Kronecker structure of the perturbed pencils $A - \lambda B + \epsilon(E_A - \lambda E_B)$, $A - \lambda B + \epsilon(Z_A - \lambda Z_B)$, and $A - \lambda B + \epsilon(T_A - \lambda T_B)$, for $\epsilon = 10^{-16}, 10^{-15}, \ldots, 10^0$. We investigate how far we can move in the tangent and normal directions before GUPTRI reports the generic Kronecker structure.

The procedure has been repeated for all cases and for 100 random perturbations (E_A, E_B), where $\|(E_A, E_B)\|_F = 1$ and $\|E_A\|_F = \|E_B\|_F$. The entries of (E_A, E_B) are uniformly distributed in $(-0.5, 0.5)$. For each case and for each perturbation $E_A - \lambda E_B$ we record the size of ϵ when GUPTRI reports the generic Kronecker structure. In Table 5 we display the smallest, median, and maximum values of ϵ for the 100 random perturbations.

Entries marked + in Table 5, represent that the generic structure was not found for any size of the perturbations. All these results were for perturbations in $\tan(A - \lambda B)$, and they indicate that for these Kronecker structures there are no or only small curvatures in the orbit at this point (pencil). Here the tangent directions are very close to orbit$(A - \lambda B)$.

Notably, the results for the perturbations $\epsilon(E_A - \lambda E_B)$ are, except for one case, similar to the results for $\epsilon(Z_A - \lambda Z_B)$. This is natural since the perturbation $E_A - \lambda E_B$ implies a translation both in the tangent space as well as the normal space directions. The structure changes appear more rapidly in the normal space, i.e., for smaller ϵ. Our computational results extend the cone example in Section 1.3 to 2-by-3 matrix pencils.

Why is the smallest perturbation $10^{-16}(Z_A - \lambda Z_B)$ enough to find the generic structure for the three cases $L_0 \oplus 2J_1(0), L_0 \oplus 2N_1$ and $3L_0 \oplus 2L_0^T$? The explanation is connected to the procedure for determining the numerical rank of matrices.

GUPTRI has two input parameters, EPSU and GAP, which are used to make rank decisions in order to determine the Kronecker structure of an input pencil $A - \lambda B$. Inside GUPTRI the absolute tolerances $\texttt{EPSUA} = \|A\|_E \cdot \texttt{EPSU}$ and $\texttt{EPSUB} = \|B\|_E \cdot \texttt{EPSU}$ are used in all rank decisions, where the matrices A and B, respectively, are involved. Suppose the singular values of A are computed in increasing order, i.e., $0 \le \sigma_1 \le \sigma_2 \le \ldots \le \sigma_k \le \sigma_{k+1} \le \ldots$;

Table 5: How far we can move in tangent and normal directions before non-generic 2-by-3 matrix pencils turn generic.

		$\epsilon(Z_A - \lambda Z_B)$			$\epsilon(T_A - \lambda T_B)$		
$A - \lambda B$	$\mathrm{cod}(A - \lambda B)$	$\epsilon_{\min}$	ϵ_{median}	$\epsilon_{\max}$	$\epsilon_{\min}$	ϵ_{median}	$\epsilon_{\max}$
$L_1 \oplus J_1(0)$	2	10^{-4}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	10^{-1}
$L_1 \oplus N_1$	2	10^{-4}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	10^{0}
$L_0 \oplus J_2(0)$	4	10^{-4}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	10^{0}
$L_0 \oplus N_2$	4	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	10^{-1}
$L_0 \oplus J_1(0) \oplus N_1$	4	10^{-4}	10^{-4}	10^{-2}	10^{-2}	10^{-1}	10^{0}
$L_0 \oplus L_1 \oplus L_0^T$	5	10^{-4}	10^{-4}	10^{-2}	10^{-2}	10^{-1}	10^{0}
$L_0 \oplus 2J_1(0)$	6	10^{-16}	10^{-16}	10^{-16}	+	+	+
$L_0 \oplus 2N_1$	6	10^{-16}	10^{-16}	10^{-16}	+	+	+
$2L_0 \oplus L_1^T$	6	10^{-4}	10^{-4}	10^{-2}	+	+	+
$2L_0 \oplus J_1(0) \oplus L_0^T$	8	10^{-5}	10^{-4}	10^{-1}	+	+	+
$2L_0 \oplus N_1 \oplus L_0^T$	8	10^{-4}	10^{-4}	10^{-3}	+	+	+
$3L_0 \oplus 2L_0^T$	12	10^{-16}	10^{-16}	10^{-16}	+	+	+

then all singular values $\sigma_k <$ `EPSUA` are interpreted as zeros. The rank decision is made more robust in practice: if $\sigma_k <$ `EPSUA` but $\sigma_{k+1} \geq$ `EPSUA`, `GUPTRI` insists on a gap between the two singular values such that $\sigma_{k+1}/\sigma_k \geq$ `GAP`. If $\sigma_{k+1}/\sigma_k <$ `GAP`, σ_{k+1} is also treated as zero. This process is repeated until an appreciable gap between the zero and non–zero singular values is obtained. In all of our tests we have used `EPSU` $= 10^{-8}$ and `GAP` $= 1000.0$.

For the most non-generic case $3L_0 \oplus 2L_0^T$, both the A-part and the B-part are zero matrices giving `EPSUA` $=$ `EPSUB` $= 0$, which in turn lead to the decision that a full rank perturbation $E_A - \lambda E_B$ times a very small ϵ is interpreted as a generic pencil. For the other two cases, either the A-part or the B-part is full rank and the other part is a zero matrix, which accordingly is interpreted to have full rank already for the smallest perturbation.

6.2.2 Versal deformations and minimal perturbations for changing a non-generic structure

In the following we illustrate how versal deformations are useful in the understanding of the relations between the different structures, by looking at requirements on perturbations to (A, B) for changing the Kronecker structure. Assume that we have the following matrix pencil with the Kronecker structure $L_1 \oplus J_1(0)$:

$$A - \lambda B = \begin{bmatrix} -\epsilon_1\lambda & \epsilon_2 & 0 \\ 0 & 0 & -\epsilon_3\lambda \end{bmatrix} \quad \text{and} \quad Z_A - \lambda Z_B = \begin{bmatrix} 0 & 0 & 0 \\ p_1 & 0 & p_2 \end{bmatrix}. \tag{6.2}$$

It was shown in [14] that $L_1 \oplus J_1(0)$ with codimension 2 is in the closure of $\mathrm{orbit}(L_1 \oplus J_1(\gamma))$ ($\gamma \neq \{0, \infty\}$ but otherwise unspecified) with codimension 1, which in turn is in the closure of $\mathrm{orbit}(L_2)$ (the generic KCF) with codimension 0. Notice in Table 4, since γ is assumed specified, $L_1 \oplus J_1(\gamma)$ has two parameters (and codimension = 2). In the discussion that follows we assume that γ is finite, non-zero but unspecified.

We will now, for this example, illustrate how perturbations in the normal space directions can be used to find more generic Kronecker structures (going upwards in the Kronecker structure hierarchy), and how we can perturb the elements in $A - \lambda B$ to find less generic matrix pencils. Since the space spanned by $Z_A - \lambda Z_B$ is the normal space, we must always first hit a more generic pencil when we move infinitesimally in normal space directions.

The KCF remains unchanged as long as $p_1 = p_2 = 0$, but for $p_1 = 0$ and $p_2 \neq 0$, the KCF is changed into $L_1 \oplus J_1(\gamma)$ (with $\gamma = p_2$). That is, by adding a component in a normal space direction, we find a more generic pencil in the closure hierarchy. Notably, the size of the required perturbation is equal to the smallest size of an eigenvalue to be interpreted as non-zero. By choosing p_1 non-zero (and p_2 arbitrary), the resulting pencil will be generic with the KCF L_2.

To find a less generic structure, we may proceed in one of the following ways:

1. Find a less generic structure in the closure of orbit$(L_1 \oplus J_1(0))$.

2. Go upwards in the closure hierarchy, to a more generic structure and then look in that orbit's closure for a less generic structure.

We know from the investigation in [14] that all structures with higher codimension than $A - \lambda B = L_1 \oplus J_1(0)$ include an L_0 block in their Kronecker structures, which in turn imply that A and B must have a common column nullspace of at least dimension 1. Therefore, the smallest perturbation that turns $L_1 \oplus J_1(0)$ less generic is the smallest perturbation that reduces the rank of

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} 0 & \epsilon_2 & 0 \\ 0 & 0 & 0 \\ \epsilon_1 & 0 & 0 \\ 0 & 0 & \epsilon_3 \end{bmatrix}.$$

The size of the smallest rank-reducing perturbation is equal to the smallest of the singular values ϵ_1, ϵ_2, and ϵ_3. By just deleting one ϵ_i, the corresponding perturbed pencil is a less generic pencil within the closure of orbit$(L_1 \oplus J_1(0))$. These three cases correspond to approach 1 above. We summarize these perturbations and the perturbations in the normal space in Table 6. Notice that approach 2 will always require a perturbation larger than $\min\{\epsilon_i\}$.

Which of the non-generic structures displayed in Table 6 is obtained by the smallest perturbation to $L_1 \oplus J_1(0)$? Mathematically, it is easy to see that the perturbations in the normal space always can be made smaller than a rank-reducing perturbation ϵ_i, since p_1 and p_2 are parameters that can be chosen arbitrary small, e.g. smaller than $\min\{\epsilon_i\}$.

However, in finite precision arithmetic, it is not clear that the smallest perturbation required to find another structure is in the normal direction. This can be illustrated by using `GUPTRI` to compute the Kronecker structures for $A - \lambda B$ as in (6.2) and perturbed as in Table 6. For `EPSU` $= 10^{-8}$, $\epsilon_2 = 1$ and $\epsilon_1 = \epsilon_3 = 10^{-10}$, `GUPTRI` uses different tolerances `EPSUA` $= 10^{-8}$ and `EPSUB` $= 10^{-18}$ for making rank decisions in A and B, respectively. It follows that for p_1 and p_2 of order 10^{-6}, `GUPTRI` still computes the Kronecker structure $L_1 \oplus J_1(0)$. However, if $p_1 = p_2 = 0$ and the B-part of the pencil is perturbed by ϵ_1 or ϵ_3, `GUPTRI` computes the less generic structures, just as shown in Table 6.

Table 6: Perturbing $A - \lambda B$ (defined in 6.2) yields the pencil $\tilde{A} - \lambda\tilde{B}$ with more or less generic structures. The codimension of the original orbit is 2.

$\|(\Delta A, \Delta B)\|_F$	$\tilde{A} - \lambda\tilde{B}$	KCF	$\mathrm{cod}(\tilde{A} - \lambda\tilde{B})$
p_1	$\begin{bmatrix} -\epsilon_1\lambda & \epsilon_2 & 0 \\ p_1 & 0 & -\epsilon_3\lambda \end{bmatrix}$	L_2	0
p_2	$\begin{bmatrix} -\epsilon_1\lambda & \epsilon_2 & 0 \\ 0 & 0 & p_2 - \epsilon_3\lambda \end{bmatrix}$	$L_1 \oplus J_1(p_2)$	1 (2)
ϵ_1	$\begin{bmatrix} -0 & \epsilon_2 & 0 \\ 0 & 0 & -\epsilon_3\lambda \end{bmatrix}$	$L_0 \oplus J_1(0) \oplus N_1$	4
ϵ_3	$\begin{bmatrix} -\epsilon_1\lambda & \epsilon_2 & 0 \\ 0 & 0 & -0 \end{bmatrix}$	$L_0 \oplus L_1 \oplus L_0^T$	5
ϵ_2	$\begin{bmatrix} -\epsilon_1\lambda & 0 & 0 \\ 0 & 0 & -\epsilon_3\lambda \end{bmatrix}$	$L_0 \oplus 2J_1(0)$	6

7 Conclusions

In this paper, we have obtained not only versal deformations for deformations of Kronecker canonical forms, but more importantly for our purposes, metrical information for the perturbation theory of matrix pencils relevant to the Kronecker canonical form. In Part II of this paper, we will explore the stratification theory of matrix pencils with the goal of making algorithmic use of the lattice of orbits under the closure relationship.

Acknowledgements

We would like to thank Jim Demmel for conveying the message that the geometry of matrix and matrix pencil space influences perturbation theory and numerical algorithms. We further thank him for many helpful discussions. In addition, the first author would like to thank the second two authors for their kind hospitality and support at Umeå University where this work and fruitful collaboration began.

References

[1] V. I. Arnold. On Matrices Depending on Parameters. *Russian Math. Surveys*, 26:29–43, 1971.

[2] 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.

[3] J.M. Berg and H.G. Kwatny. A canonical parameterization of the Kronecker form of a matrix pencil. Preprint, 1994. To appear in *Automatica.*

[4] T. Brocker and L. Lander. *Differential Germs and Catastrophes.* Cambridge University Press, 1975.

[5] J.W. Bruce and P.J. Giblin. *Curves and Singularities.* Cambridge University Press, 1991.

[6] S.-N. Chow, C. Li, and D. Wang. *Normal Forms and Bifurcation of Planar Vector Fields.* Cambridge University Press, 1994.

[7] J. Demmel. On structured singular values. In *Proceedings of the 27th Conference on Decision and Control*, Austin, TX, Dec 1988. IEEE.

[8] 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.*

[9] J. Demmel and B. Kågström. Computing stable eigendecompositions of matrix pencils. *Lin. Alg. Appl.*, 88/89:139–186, April 1987.

[10] J. Demmel and B. Kågström. Accurate solutions of ill-posed problems in control theory. *SIAM J. Mat. Anal. Appl.*, 9(1):126–145, January 1988.

[11] 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.

[12] 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.

[13] A. Edelman and H. Murakami. Polynomial Roots from Companion Matrix Eigenvalues. *Mathematics of Computations, to appear*, 1995.

[14] 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.*

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

[16] G. Golub and C. Van Loan. *Matrix Computations.* Second Edition. Johns Hopkins University Press, Baltimore, MD, 1989.

[17] M. Golubitsky and V. Guillemin. *Stable Mappings and their Singularities.* Springer Verlag, 1973.

[18] U. Helmke and D. Hinrichsen. Canonical forms and orbit spaces of linear systems. *IMA J. Mathematical Control & Information*, pages 167–184, 1986.

[19] D. Hinrichsen and D. Prätzel-Wolters. A wild quiver in linear systems theory. *Linear Algebra Applications*, pages 143–175, 1987.

[20] D. Hinrichsen and D. Prätzel-Wolters. A Jordan canonical form for reachable linear systems. *Linear Algebra Applications*, pages 489–524, 1989.

[21] 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.

[22] 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.

[23] V. Kublanovskaya. AB-algorithm and its modifications for the spectral problem of linear pencils of matrices. *Num. Math.*, 43:329–342, 1984.

[24] Y.-C. Lu. *Singularity Theory and an Introduction to Catastrophe Theory.* Springer Verlag, 1976.

[25] D. Prätzel-Wolters. Canonical forms for linear systems. *Linear Algebra Applications*, pages 437–473, 1983.

[26] M. Shub and S. Smale. Complexity of Bezout's theorem II: volumes and probabilities. In F. Eyssette and A. Galligo, editors, *Computational Algebraic Geometry*, Progress in Mathematics Vol. 109, pages 267–285. Birkhauser, 1993.

[27] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory.* Academic Press, New York, 1990.

[28] L. Stolovitch. On the computation of a versal family of matrices. *Numerical Algorithms*, 4:25–56, 1993.

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

[30] P. Van Dooren. The generalized eigenstructure problem in linear system theory. *IEEE Trans. Autom. Contr.*, AC-26(1):111–129, 1981.

[31] P. Van Dooren. Reducing subspaces: Definitions, properties and algorithms. In B. Kågström and A. Ruhe, editors, *Matrix Pencils*, pages 58–73. Springer-Verlag, Berlin, 1983. Lecture Notes in Mathematics, vol. 973, Proceedings, Pite Havsbad, 1982.

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

[33] J. H. Wilkinson. *The Algebraic Eigenvalue Problem.* Oxford Science Publications, 1965.

[34] 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.