ABSTRACT

OKAY, IRFAN. The Additional Dynamics of the Least Squares Completions of Linear Differential Algebraic Equations. (Under the direction of Dr. Stephen L. Campbell.)

Differential equations of the form $F(x', x, t) = 0$ with $F_{x'}$ singular arise naturally in many applications and are generally called differential algebraic equations (DAE). There has been an extensive amount of research on numerical solutions of DAEs in recent years. While the classical ODE methods such as backward differentiation and Runga-Kutta methods can be used to numerically solve DAEs, they require the problem to have lower index or special structure.

One approach proposed for solving more general, higher index DAEs is called explicit integration (EI). The original DAE is differentiated a number of times based on certain parameters and the new system of equations is solved using nonlinear least squares methods. The result is a computed ODE whose solutions contain the solutions of the DAE. It is called the least squares completion (LSC). This ODE is then numerically integrated by a classical numerical method.

The EI method is computationally efficient and can be applied to a wide class of DAEs. However, the dynamics of the additional solutions present in the completion can effect the numerical integration, sometimes causing the numerical solutions to move away from the solution manifold. In this thesis, we analyze the additional dynamics of LSCs for linear DAEs. Starting with linear constant coefficient systems, we first examine the structure of the additional dynamics created by the standard LSC and then introduce two methods to modify the completion process so that the LSC will have additional dynamics with desired stability characteristics. The rate of stabilized convergence can be determined a priori by substituting an appropriate value for a parameter. We then extend the results to linear time variable systems.

The Additional Dynamics of the Least Squares Completions of Linear Differential Algebraic Equations

by
Irfan Okay

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Mathematics

Raleigh, North Carolina

2008

APPROVED BY:

Dr. Negash Medhin	Dr. Hien Tran
Dr. Ernie Stitzinger	Dr. Stephen L. Campbell Chair of Advisory Committee

To my Parents . . .

BIOGRAPHY

Irfan Okay was born in Mardin, Turkey in 1978. He obtained his undergraduate degree from Mersin University, Department of Mathematics. He entered the graduate program in Mathematics at NC State University in the fall of 2004.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor Dr. Stephen L. Campbell. This work would not have been possible without his invaluable advice and generous support. His extraordinary knowledge and understanding of mathematics have benefited me greatly. It has been a privilege to work with him.

I would like to express my sincere thanks to Dr. Ernie Stitzinger for all the prompt assistance he has provided over the years as a graduate administrator. I greatly appreciate it. Special thanks to Dr. Negash Medhin for all his help and assistance as a committee member. I also have to thank my committee members Dr. Hien Tran, Dr. Zhilin Li , Dr. Kazufumi Ito, and our graduate secretary Denise Seabrooks, who were of great help. Finally, I owe many thanks to Dr. Peter Kunkel for his valuable contributions which added greatly to the thesis.

TABLE OF CONTENTS

Chapter 1

Introduction

1.1 Outline of the Thesis

Chapter 1 surveys the general DAE background and literature. The fundamental concepts such as solvability and index are introduced. The last section describes the least squares completion (LSC) method that we will analyze in the thesis and gives some numerical examples to illustrate the process.

In Chapter 2-4 we study linear time invariant DAEs. In Chapter 2 we analyze the additional dynamics of LSC defined by the standard derivative array. Using canonical decomposition, we first identify the part of the DAE that creates the additional dynamics. We then form the derivative array and solve the equations using linear algebra to obtain an analytical formula for the completion. The eigenstructure of the completion is analyzed to determine the nature of the additional dynamics.

In Chapter 3 and Chapter 4 we introduce two new methods to obtain LSCs with desired additional dynamics. They are called stabilized LSC and alternative stabilized completion. The stabilized LSC is based on forming the derivative array using stabilized differentiation while the alternative stabilized completion uses an index one formulation of the DAE to obtain a completion. We analyze the stability of the completions and discuss some of the numerical issues.

In Chapter 5 and 6, we apply these two techniques to linear time varying systems. For the extension of the stabilized LSC we use a a more general technique that combines the machinery developed in both Chapter 3 and Chapter 4. This enables us to determine the behavior of additional dynamics without obtaining an explicit formula for the completion which would be difficult for time variable systems. The extension of the alternative stabilized completion to LTV systems is more straightforward. Certain computational issues are also discussed and a comparison between two techniques is given.

The last chapter summarizes the results we have obtained and discusses several possible future research topics. At the end of the chapter is a list of presentations and papers where this research has appeared.

1.2 DAE Basics

A system of differential equations of the form

$$F(x', x, t) = 0 \tag{1.1}$$

where $F_{x'} = \partial F/\partial x'$ is identically singular and $x' = dx/dt$, is called a differential-algebraic equation(DAE). DAEs have become increasingly important in recent years as many physical processes can be easily modeled as a nonlinear implicit system of DAEs. Some of the well known examples include trajectory prescribed path control, systems of rigid bodies, problems in constrained mechanics, electrical networks and chemical reactions [9]. A classical example of a DAE arising from mechanical systems is the equation describing the motion of a pendulum.

$$x'' = \lambda x \tag{1.2a}$$

$$y'' = \lambda y - g \tag{1.2b}$$

$$0 = x^2 + y^2 - L^2 \tag{1.2c}$$

Here g is the gravitational constant, L is the length of the pendulum, (x, y) are the coordinates of the ball of the infinitesimal mass attached at the end of the pendulum and λ denotes an unknown function corresponding to a Lagrange multiplier and λx is force. The Euler-Lagrange formulation [39] of many problems in constrained mechanics give rise to DAEs. Note that if we introduce the velocity variables $v_x = x'$ and $v_y = y'$, then (1.2) takes the form

$$
\begin{aligned}
x' &= v_x && \text{(1.3a)}\\
y' &= v_y && \text{(1.3b)}\\
v_x' &= -\lambda x && \text{(1.3c)}\\
v_y' &= -\lambda y - g && \text{(1.3d)}\\
0 &= x^2 + y^2 - L^2 && \text{(1.3e)}
\end{aligned}
$$

which is now in the standard form (1.1).

The variables x, y, v_x, v_y are called differential variables since their derivatives appear in (1.3) and λ is called an algebraic variable since λ' does not appear in (1.3). In many cases algebraic and differential variables are intertwined in a complex manner rendering the equations inextricable by algebraic manipulations. Because of the singularity condition on $F_{x'}$, DAEs always contain pure algebraic equations called constraints. For example the equation (1.3e) is a constraint. However, not all the constraints are given explicitly. DAEs can also contain constraints that are revealed only after differentiating explicitly given equations. They are called hidden constraints. For example, differentiating (1.3e) once we get

$$xx' + yy' = 0 \tag{1.4}$$

Then, substituting (1.3a) and (1.3b) we obtain the hidden constraint

$$xv_x + yv_y = 0 \tag{1.5}$$

Working with DAEs presents analytical and numerical difficulties that are not present when working with ODEs [9], [3]. For example the solutions of a DAE may not be equally smooth in all components. In general, the differential variables will be the smoother than the algebraic variables. This is because integration is being used to calculate a differential variable, which a smoothing process, while differentiation is used to reveal hidden algebraic variables and each differentiation reduces the degree of smoothness by one.

Because of the existence of algebraic constraints, the solutions of a DAE form a manifold called the solution manifold. Only the initial conditions that lie on the solution manifold accept a solution to the DAE. These are called consistent initial conditions. A particular solution of the DAE is thus a curve moving on this manifold. Under certain conditions a DAE can be thought of as an ODE defined on the solution manifold [57].

There has been extensive research on solving (1.1) numerically. The ODE methods such as backward differentiation and Runge Kutta methods can be applied to DAE's [9], [41], [51]. However, they are only suitable for lower index problems (to be defined) and requires the problem to have a specific structure.

One general method for solving (1.1) numerically is to find an ODE whose solutions contains the solutions of the DAE and integrate the ODE using classical ODE methods. An ODE that contains the solutions of the DAE is called a completion. This can be done by differentiating the original equations until the larger systems of equations can define an ODE. The minimum number of differentiations needed to differentiate the DAE or part of the DAE to obtain such an ODE is called the index of the DAE [21].

In our example, differentiating (1.5) we obtain

$$xv_x' + yv_y' + v_x^2 + v_y^2 = 0 \tag{1.6}$$

Then substituting v_x', v_y' from (1.3), we get

$$\lambda = \frac{1}{L^2}(v_x^2 + v_y^2 - yg) \tag{1.7}$$

Now substituting λ into (1.3c) and (1.3d), and differentiating the equation for λ we arrive at the ODE

$$\begin{aligned}
x' &= v_x && (1.8a)\\
y' &= v_y && (1.8b)\\
v_x' &= -(v_x^2 + v_y^2 - v_y g)x && (1.8c)\\
v_y' &= -(v_x^2 + v_y^2 - v_y g)y - g && (1.8d)\\
\lambda' &= \frac{1}{L^2}(v_x^2 + v_y^2 - yg)' && (1.8e)
\end{aligned}$$

Since we had to differentiate the constraint equation (1.3e) three times, the index is 3. The DAE is called higher index if the index is bigger than one. The index is in some sense an indication of complexity of the DAE or more precisely how close it is to being an ODE. So, for example, according to this definition, an ODE has index zero, while a pure algebraic system will have index one. What we have called the index is more accurately called the differentiation index. There are other type of indices but we do not need them in this thesis.

We should note that a completion obtained this way is not unique [19]. It depends on the equations used and the solution method. For example, a pure algebraic DAE $x = t$ can be differentiated once to obtain the completion $x' = 1$. However, we can also add this completion to the original equation to obtain $x' + x = t + 1$, which is a different completion of $x = t$.

Among some other techniques for numerically solving general DEAs are index-one integration [42], [43], [46], and coordinate partitioning methods [1].

1.3 Solvability

Intuitively, a solution of (1.1) on an interval I is a continuously differentiable function $y(t)$ satisfying

$$F(y'(t), y(t), t) = 0$$

for all $t \in I$. However, DAEs can exhibit aberrant characteristics in general when it comes to the structure of solutions. Therefore we need a more precise definition of solvability. The following definition, which is referred to as geometric solvability, will suffice for our purposes. More technical definitions and discussions on solvability can be found in [22].

Definition 1 *Let I be an open subset of R, Ω a connected open subset of R^{2s+1}, and F a differentiable function from Ω to R. The DAE is solvable on I in Ω if there is an r-dimensional family of solutions $\phi(t, c)$ defined on a connected open set $I \times \tilde{\Omega}$, $\tilde{\Omega} \subseteq R^r$, such that*

- *$\phi(t, c)$ is defined on all of I for each $c \in \tilde{\Omega}$*

- *$(\phi_t(t,c), \phi(t,c), t) \in \Omega$ for $(t,c) \in I \times \tilde{\Omega}$.*

- *If $\psi(t)$ is any solution with $(\psi'(t), \psi(t), t) \in \Omega$, then $\psi(t) = \phi(t,c)$ for some $c \in \tilde{\Omega}$*

- *The graph of ϕ as a function of (t,c) is an (r+1)-dimensional manifold.*

Basically, the definition tells us that the DAE locally has a unique solution manifold and each solution is uniquely determined by the initial condition. Existing numerical methods either require solutions to exist or the solution manifold to have a specific structure to work. Some existence results have been obtained using differential geometric techniques [58], [52], [53], [55], [54], [56]. In this section, we will give a characterization of solvability that is also computationally verifiable [22]. For nonlinear systems only sufficient conditions can be expected in general.

Suppose we differentiate (1.1) k times with respect to t. Then, we get the extended system of equations

$$F = 0 \tag{1.9a}$$
$$\frac{d}{dt}F = 0 \tag{1.9b}$$
$$\vdots \tag{1.9c}$$
$$\frac{d^k}{dt^k}F = 0 \tag{1.9d}$$

which is called a derivative array and denoted by

$$G = G(x', w, x, t) = 0 \tag{1.10}$$

where

$$w = [x^{(2)}, \ldots, x^{(k+1)}]. \tag{1.11}$$

Definition 2 *A system of algebraic equations*

$$A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = b$$

is called 1-full *with respect to x_1 if x_1 is uniquely determined for any consistent b [13].*

Now suppose that the following assumptions are satisfied for both k and $k+1$ in some neighborhood:

(A1) Sufficient smoothness of $G = 0$.

(A2) $G = 0$ is consistent as an algebraic equation.

(A3) $J = [\ G_{x'} \quad G_w\ ]$ is 1-full and has constant rank independent of (x', w, x, t).

(A4) $[\ G_{x'} \quad G_w \quad G_x\ ]$ has full row rank independent of (x', w, x, t).

Given the above definitions and assumptions we have;

Theorem 1 *[22], [40] Suppose that the derivative array $G(x', w, x, t)$ satisfies the conditions (A1)–(A4) in a neighborhood. Then, the DAE (1.1) is geometrically solvable with the solution manifold S_k, where*

$$S_k = \{(t, x) | G(\overline{v}, \overline{w}, x, t) = 0, \textit{for some } (\overline{v}, \overline{w})\}$$

For linear systems, (A1)–(A4) is almost equivalent to solvability [16]. To illustrate how these conditions relate to the geometric solvability, consider the linear time variable DAE

$$A(t)x' + B(t)x = f(t) \tag{1.12}$$

Differentiating (1.12) k times with respect to t gives us the derivative array

$$J \begin{bmatrix} x' \\ w \end{bmatrix} = -\mathcal{F}x + g. \tag{1.13}$$

where

$$J = \begin{bmatrix} A & 0 & 0 & \cdots & 0 \\ A' + B & A & 0 & \cdots & 0 \\ A'' + 2B' & 2A' + B & A & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ . & . & . & \cdots & . \end{bmatrix}, \mathcal{F} = \begin{bmatrix} B \\ B' \\ B'' \\ \vdots \\ B^{(k)} \end{bmatrix}, g = \begin{bmatrix} f \\ f' \\ f'' \\ \vdots \\ f^{(k)} \end{bmatrix}$$

Suppose that assumptions (A1)–(A4) holds for (1.13). Note that we have $J = [\ G_{x'} \quad G_w\ ]$. Therefore, by smoothness and 1-fullness, there exists a nonsingular smooth D such that

$$DJ = \begin{bmatrix} I_n & 0 \\ 0 & C \end{bmatrix} \tag{1.14}$$

Then, multiplying (1.13) by D we get

$$\begin{bmatrix} I_n & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} x' \\ w \end{bmatrix} = -D\mathcal{F}x + Dg. \tag{1.15}$$

Since the right hand side is a function of just x and t, the first block row gives us an ODE

$$x' = V^T(-D\mathcal{F}x + Dg) = h(x,t) \tag{1.16}$$

where $V = \begin{bmatrix} I_n & 0 & \cdots & 0 \end{bmatrix}^T$. On the other hand, by smoothness and constant rank assumptions, there exists a smooth matrix function Z of maximal rank satisfying $ZJ = 0$. Then, together with the assumption (A4), this implies that

$$\text{rank}(Z) = \text{rank}(Z\mathcal{F}) \tag{1.17}$$

Therefore, the equation

$$0 = Z(-\mathcal{F}x + g) \tag{1.18}$$

comprises all the constraints of the original DAE. In other words, (1.18) precisely defines the solution manifold. The uniqueness of the manifold follows from the maximality of Z. Now, the solutions of the DAE are given by the solutions of the ODE with the initial conditions defined by (1.18). Since the solutions of the ODE are uniquely determined by the initial conditions, the solutions of the DAE are therefore uniquely determined by the consistent initial conditions.

The above definition of solvability is also vital in that when computing a completion of a DAE, it guarantees that the solutions of the completion will coincides with those of the DAE on the manifold.

What is possible to prove or compute depends on the structure of the DAE. Here are some of the important classes of DAE's:

- Linear time invariant DAE (LTI)

$$Ax' + Bx = f(t) \tag{1.19}$$

where A and B are constant and A is singular. This is one of the most basic and well understood classes of DAEs. They have been studied extensively [10], [11]. Besides their important applications, they are also an ideal class of problems to test and develop methods intended for more general classes of DAEs. The matrix pencil $\lambda A + B$ is called regular if $\det(\lambda A + B)$ is not identically zero as a function of λ. The solvability of (1.19) is equivalent to the regularity of the pencil [36].

To illustrate this relationship, suppose that we want to solve (1.19) numerically using implicit Euler starting at t_0 with constant step-size h. Let $t_n = t_0 + nh$, x_n be the estimate for $x(t_n)$ and $c_n = c(t_n)$ for $c = f, A, B$. Then, applying the implicit Euler to (1.19) we obtain

$$A\frac{x_n - x_{n-1}}{h} + Bx_n = f_n$$

which becomes

$$(A + hB)x_n = Ax_{n-1} + hf_n$$

and $A + hB$ has to be regular in order to uniquely determine x_n given x_{n-1}. Therefore, the matrix pencil $A + \lambda B$ needs to be regular.

- Linear time variable DAE (LTV)

$$A(t)x' + B(t)x = f(t) \tag{1.20}$$

where $A(t)$ and $B(t)$ are matrix functions of t, and $A(t)$ is singular possibly for all t. Similar to the constant coefficient case, the DAE is called regular if $\det(A(t) + \lambda B(t))$ is not zero as a function of λ for all t. However, the regularity is no longer equivalent to solvability for linear time varying systems [9]. It actually turns out to be quite independent.

The initial work on LTV systems was based on the standard canonical form [47], [59], [12], [30], [50]. Later a numerical method was introduced that was based on a more general canonical form which covers all solvable systems [13], [14], [15], [16]. Some applications of the numerical method can be found in [29], [31].

LTV DAEs share important structural similarities with nonlinear systems while still benefiting from the linearity. Therefore, understanding the linear time variable DAEs is a significant milestone towards the understanding of nonlinear systems.

- Semi-explicit index-1 DAE

$$x' = f(x, y, t) \tag{1.21a}$$

$$0 = g(x, y, t) \tag{1.21b}$$

where g_y is nonsingular. Note that if we differentiate the constraint equation (1.21b) once we get

$$x' = f(x, y, t) \tag{1.22}$$

$$g_x(x, y, t)x' + g_y(x, y, t)y' = 0 \tag{1.23}$$

Since g_y is nonsingular, the system (1.23) is an ODE, therefore (1.21) has index one. A semi-explicit index one DAE is also called a Hessenberg index-1 DAE.

- Hessenberg index-2 DAE

$$x' = f(x, y, t) \tag{1.24a}$$

$$0 = g(x, , t) \tag{1.24b}$$

where $(dg/dx)(df/dy)$ is nonsingular.

- Hessenberg index-3 DAE

$$x' = f(x, y, z, t) \tag{1.25a}$$

$$y' = g(x, y, t) \tag{1.25b}$$

$$0 = h(y, t) \tag{1.25c}$$

where $(dh/dy)(dg/dx)(df/dz)$ is nonsingular. The example (1.3) is a Hessenberg index-3 DAE. Many mechanical systems fall in this category. Hessenberg systems are solvable [18], [41].

1.4 Least Squares Completions

In this section we will describe the least squares completion which is the basis of the explicit integration process. We have already noted that given a derivative array $G(v, w, x, t)$ satisfying (A1)–(A4), there are usually many ways one can obtain an ODE [19]. However, in order to develop efficient numerical methods that can be applied to a wide class of DAEs, we first need an algorithm that will always produce an ODE given a sufficiently large derivative array. The ODE should also have good smoothness properties even though it is computed numerically pointwise.

For the nonlinear system (1.10), let $H(v, w) = G(v, w, x, t)$ for a given (x, t). Given an initial guess (v_0, w_0), we shall solve (1.10) for (v, w) numerically using the generalized Gauss-Newton iteration [8]

$$[v_{n+1}, w_{n+1}] = [v_n, w_n] - [H_v(v_n, w_n), H_w(v_n, w_n)]^{\dagger} H(v_n, w_n), \tag{1.26}$$

where $A^{\dagger}b$ is the minimum norm least squares solution of $Ax = b$ [24]. Modifications of (1.26) and other computational issues are discussed in [25], [27]. It is important to note that (1.26) is done for each possible (x, t) so that both the terms on the right hand side of (1.26) depend on x, t.

In [17] it is shown that under the assumptions (A1)–(A4) that if (v_0, w_0) is close enough to values for a solution of (1.1) and (x, t) is close enough to being consistent, then the iteration (1.26) converges. Let (v^*, w^*) be the limit of the iteration (1.26). Then the limit (v^*, w^*) satisfies the *least squares equation* (LSE)

$$[H_v(v^*, w^*), H_w(v^*, w^*)]^T H(v^*, w^*) = 0 \tag{1.27}$$

Note that (1.27) is not equivalent to (1.10) since $[H_v \; H_w]$ does not have full row rank. We have used H to simplify our notation but the least squares equations are actually

$$[G_v(v^*, w^*, x, t), G_w(v^*, w^*, x, t)]^T G(v^*, w^*, x, t) = 0 \tag{1.28}$$

Theorem 2 *[20] Suppose that the derivative array $G(v, w, x, t)$ satisfies the conditions (A1)–(A4) on an open neighborhood of a consistent initial condition (v_0, w_0, y_0, t_0) and let*

$$\tilde{G}(v, w, x, t) = [G_v \; G_w]^T G(v, w, x, t)$$

Then, $\tilde{G} = 0$ determines locally a unique h such that

$$v = x' = h(x, t) \tag{1.29}$$

and (v_0, x_0, t_0) lies on the graph of h. Moreover, the degree of smoothness of h in (x, t) is at most one order less than that of G in (x, t).

Some computational aspects of the method are studied in [32], [33]. Lets again look at the linear system

$$A(t)x' + B(t)x = f(t) \tag{1.30}$$

The analogue of (1.27) for this system will be

$$J^T J \begin{bmatrix} x' \\ w \end{bmatrix} = J^T(-\mathcal{F}x + g) \tag{1.31}$$

The right hand side of (1.31) is a function of just x and t. Therefore, the solutions of this system will produce $[x', w]$ in terms of x and t. Since $J^T J$ is singular, there will be generally many $[x', w]$ given (x, t). However, the theorem states that all of them will return the same x' and the dependence of x' on x, t will be continuous.

The basic idea of [17], which is called explicit integration, is to numerically compute an ODE whose solutions contain the solutions of the DAE. The LSC is the ODE used in the explicit integration process. Using near consistent initial conditions one can integrate the completion numerically by a classical method to estimate the solutions of the DAE. One difficulty is the effect of additional dynamics of the completion on the numerical integration process. If the solution manifold is not asymptotically stable inside the completion, then numerical results can move away from the manifold during the integration. While some drifting can be tolerated in certain problems, it can have serious consequences where the manifold represents important physical processes. In this thesis, we will analyze the nature of these additional dynamics and develop methods to modify the completion process in order to obtain better additional dynamics.

Existing stabilization techniques include enforcing the constraints using certain parameters [37], [7] or numerical constraint preserving techniques of [4], [5], [26], [33]. Probably the best known in applications is Baumgarte stabilization [7]. However, these methods either require the constraints to be explicitly known or involve problem specific numerical manipulations that carry a high computational cost. In this thesis, we first determine the analytical structure of the additional dynamics in a LSC, then, by incorporating the ideas of [7], [42], we modify the derivative array in a way that will produce a completion with desired additional dynamics so that the rate of stability can be determined a priori by inserting an appropriate value for a parameter λ.

Another important concept that is closely related to least squares solutions is generalized inverses (Moore-Penrose). We will frequently use some basic properties of generalized inverses when calculating the LSCs. While there are various equivalent definitions of Moore-Penrose inverse, the following definition will be sufficient for the thesis. A detailed analysis of generalized inverses can be found in [24].

Definition 3 *If $A \in \mathbb{C}^{m\times n}$, then the generalized inverse (Moore-Pnerose) of A is defined to be the unique matrix $A^\dagger$ satisfying*

1. $AA^\dagger = P_{R(A)}$
2. $A^\dagger A = P_{R(A^\dagger)}$, *where P_Z is the orthogonal projector onto Z.*

As can be seen from this definition, a generalized inverse will be the same as the ordinary inverse when the matrix is invertible. While it is a natural generalization of the ordinary inverse, we don't have $(AB)^\dagger = B^\dagger A^\dagger$ in general. However, if P is an orthogonal matrix, then we have $(PB)^\dagger = P^\dagger B^\dagger$ [24]. Using this property, one can use the singular value decomposition to calculate the generalized inverse of a matrix.

Suppose that A has the singular value decomposition

$$A = U^T \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V \tag{1.32}$$

where U and V are orthogonal and D is diagonal. Then, since the genralized inverse just coincides with the ordinary inverse for an invertible matrix, the definition implies

$$A^{\dagger} = V^{T} \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix} U \tag{1.33}$$

The next two results are some of the basic properties of generalized inverses in connection with least squares solutions and will be used frequently in the rest of the thesis.

Theorem 3 *[24] Suppose that $A \in \mathbb{C}^{m\times n}$ and $b \in \mathbb{C}^{m\times n}$. Then the following are equivalent*

1. *u is a least squares solution of $Ax = b$.*
2. *u is a solution of $Ax = AA^{\dagger}b$.*
3. *u is a solution of $A^*Ax = A^*b$, where A^* is the conjugate transpose of A.*
4. *u is of the form $A^{\dagger}b + h$ where $h \in N(A)$, the null space of A. Also, $A^{\dagger}b$ is the minimal norm least squares solution of $Ax = b$.*

Theorem 4 *[24] Suppose that we have $R(Y) \subseteq R(X)$ and $R(Y^*) \subseteq R(X^*)$, where $R(Z)$ denotes the range of Z. Then*

$$\begin{bmatrix} X & 0 \\ Y & Z \end{bmatrix}^{\dagger} = \begin{bmatrix} X^{\dagger} & 0 \\ -X^{\dagger}YZ^{\dagger} & Z^{\dagger} \end{bmatrix}$$

In practice, it may not be possible to identify which part of the DAE needs to be differentiated or how many times. While there are some techniques to determine the index of the system [35], [48] [49], it might be necessary or practical to perform extra differentiations to ensure the completeness of the derivative array. The next result is very important in that respect, as it shows that the least squares completion is not changed by extra, or reduced, differentiation. It appears in a modified form in [20].

Theorem 5 *Suppose that G is a derivative array which may have been formed by differentiating different equations in (1.1) a different number of times and with Jacobian*

J which is large enough so that assumptions (A1)–(A4) hold. Suppose that we differentiated F some additional times so that we have additional equations $\widetilde{G} = 0$. Then the least squares equations for this larger set of equations are

$$
\begin{aligned}
J^T G &= 0 &\quad (1.34a)\\
\widetilde{G} &= 0 &\quad (1.34b)
\end{aligned}
$$

Proof. The least squares equations for the larger derivative array $\widehat{G}$ are

$$
\widehat{J}^T \widehat{G} = \begin{bmatrix} J^T & J_1^T \\ 0 & J_2^T \end{bmatrix} \begin{bmatrix} G \\ \widetilde{G} \end{bmatrix} = 0. \tag{1.35}
$$

Performing a row compression of J^T we get

$$
\left[\begin{array}{c|c} R & * \\ 0 & J_3 \\ \hline 0 & J_2^T \end{array}\right] \begin{bmatrix} G \\ \widetilde{G} \end{bmatrix} = 0
$$

where R is full row rank. Now the fact that the (A1)–(A4) assumptions hold for J means that J and $\widehat{J}$ have the same co-rank since the corank equals the number of constraints defining the solution manifold. Thus the nullity of J^T and $\widehat{J}^T$ are the same and the nullity of J^T is the nullity of the matrix R. Thus the matrix $\begin{bmatrix} J_3 \\ J_2^T \end{bmatrix}$ is full column rank. Hence (1.35) is equivalent to $RG = 0$ and $\widetilde{G} = 0$ which is (1.34). □

The next two examples illustrate how to analytically calculate a LSC. Note that the process we will describe is for analysis only. The numerical calculation of LSCs is done differently, using specific numeric codes [17].

Example 1 *Consider the following linear time variable DAE*

$$
\begin{bmatrix} 0 & t \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, I = [-1, 1]
$$

Differentiating the DAE twice, we get the Jacobian

$$J_3 = \begin{bmatrix} A & 0 & 0 \\ A'+B & A & 0 \\ A''+2B' & 2A'+B & A \end{bmatrix} = \left[\begin{array}{cc|cc|cc} 0 & t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline -1 & 1 & 0 & t & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & -1 & 2 & 0 & t \\ 0 & 0 & 0 & -1 & 0 & 0 \end{array}\right]$$

with

$$g = \begin{bmatrix} f \\ f' \\ f'' \end{bmatrix} = \left[\begin{array}{c} f_1 \\ f_2 \\ \hline f_1' \\ f_2' \\ \hline f_1'' \\ f_2'' \end{array}\right], \mathcal{F} = \begin{bmatrix} B \\ B' \\ B'' \end{bmatrix} = \left[\begin{array}{cc} -1 & 0 \\ 0 & -1 \\ \hline 0 & 0 \\ 0 & 0 \\ \hline 0 & 0 \\ 0 & 0 \end{array}\right], J_2 = \left[\begin{array}{cc|cc} 0 & t & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline -1 & 2 & 0 & t \\ 0 & -1 & 0 & 0 \end{array}\right]$$

We see that corank$(J_3) =$ *corank*$(J_2) = 2$*. Thus the least squares equation is given by*

$$J_3^T J_3 w = J_3^T(-\mathcal{F}x + g)$$

which becomes

$$\left[\begin{array}{cc|cc|cc} 1 & -1 & 0 & -t & 0 & 0 \\ -1 & t^2+2 & 0 & t & 0 & 0 \\ \hline 0 & 0 & 1 & -2 & 0 & -t \\ -t & t & -2 & t^2+5 & 0 & 2t \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -t & 2t & 0 & t^2 \end{array}\right] \begin{bmatrix} x' \\ w \end{bmatrix} = \left[\begin{array}{cc|cc|cc} 0 & 0 & -1 & 0 & 0 & 0 \\ t & 0 & 1 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & t & 0 & 2 & -1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & t & 0 \end{array}\right] (-\mathcal{F}x+g)$$

Now, performing the following elementary row operations on both sides of the equation in the given order $(tR_1+R_4) \to R_4, (2R_2+R_4) \to R_4, (tR_4+R_1) \to R_1, (-tR_4+R_2) \to R_2, (R_1+R_2) \to R_2$*, we obtain*

$$\left[\begin{array}{cc|cc|cc} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & t^2+1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & -2 & 0 & -t \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -t & 2t & 0 & t^2 \end{array}\right] \begin{bmatrix} x' \\ w \end{bmatrix} = \left[\begin{array}{cc|cc|cc} 0 & 0 & -1 & 0 & 0 & -t \\ t & 0 & 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & t & 0 \end{array}\right] (-\mathcal{F}x + g)$$

The first block row of the equation then gives us

$$\begin{bmatrix} 1 & -1 \\ 0 & t^2+1 \end{bmatrix} x' = \left[\begin{array}{cc|cc|cc} 0 & 0 & -1 & 0 & 0 & -t \\ t & 0 & 0 & -1 & 0 & 0 \end{array}\right] (-\mathcal{F}x + g)$$

Substituting in for $-\mathcal{F}x + g$, we get

$$\begin{bmatrix} 1 & -1 \\ 0 & t^2+1 \end{bmatrix} x' = \left[\begin{array}{cc|cc|cc} 0 & 0 & -1 & 0 & 0 & -t \\ t & 0 & 0 & -1 & 0 & 0 \end{array}\right] \left(- \left[\begin{array}{cc} -1 & 0 \\ 0 & -1 \\ \hline 0 & 0 \\ 0 & 0 \\ \hline 0 & 0 \\ 0 & 0 \end{array}\right] x + \left[\begin{array}{c} f_1 \\ f_2 \\ \hline f_1' \\ f_2' \\ \hline f_1'' \\ f_2'' \end{array}\right] \right)$$

Thus we arrive at

$$\begin{bmatrix} 1 & -1 \\ 0 & t^2+1 \end{bmatrix} x' = - \begin{bmatrix} 0 & 0 \\ -t & 0 \end{bmatrix} x + \begin{bmatrix} -f_1' - tf_2'' \\ tf_1 - f_2' \end{bmatrix}$$

So the completion becomes

$$\begin{aligned} x' &= \begin{bmatrix} 1 & -1 \\ 0 & t^2+1 \end{bmatrix}^{-1} \left(- \begin{bmatrix} 0 & 0 \\ -t & 0 \end{bmatrix} x + \begin{bmatrix} -f_1' - tf_2'' \\ tf_1 - f_2' \end{bmatrix} \right) && (1.36) \\ &= (1/t^2+1) \left\{ \begin{bmatrix} t & 0 \\ t & 0 \end{bmatrix} x + \begin{bmatrix} (t^2+1)(-f_1' - tf_2'') + (tf_1 - f_2') \\ tf_1 - f_2' \end{bmatrix} \right\} && (1.37) \end{aligned}$$

Example 2 *Lets now consider the nonlinear DAE*

$$y' = x \tag{1.38a}$$

$$0 = y^3 - x^3 \tag{1.38b}$$

The DAE is solvable with solutions $x = ce^t, y = ce^t$. The derivative array with $k = 1$ is

$$y' - x = 0 \tag{1.39a}$$

$$y^3 - x^3 = 0 \tag{1.39b}$$

$$y'' - x' = 0 \tag{1.39c}$$

$$3y^2y' - 3x^2x' = 0 \tag{1.39d}$$

Therefore we have

$$J = [G_v, G_w] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \\ -3x^2 & 3y^2 & 0 & 0 \end{bmatrix} \tag{1.40}$$

Solutions are well defined even for $c = 0$. However, (1.40) is 1-full and constant rank only if $x \neq 0$. An easy calculation shows that from (1.39) we can calculate a completion of (1.38) as

$$y' = x \tag{1.41a}$$

$$x' = x \tag{1.41b}$$

However, the least squares equations $[G_v, G_w]^T G = 0$ are

$$y' - x' + 3x^2(3y^2y' - 3x^2x') = 0 \tag{1.42a}$$

$$(y' - x) + 3y^2(3y^2y' - 3x^2x') = 0 \tag{1.42b}$$

$$y'' - x' = 0 \tag{1.42c}$$

$$0 = 0 \tag{1.42d}$$

Assuming that x, y are nonzero, (1.42) reduces to

$$3y^2y' - 3x^2x' = 0 \tag{1.43a}$$
$$y' - x = 0 \tag{1.43b}$$
$$y'' - x' = 0 \tag{1.43c}$$

which gives us the nonlinear LSC

$$x' = y^2x^{-1} \tag{1.44a}$$
$$y' = x \tag{1.44b}$$

which is not defined if $x = 0$*. Note that as a set of equations, (1.38) is equivalent to*

$$y' = x \tag{1.45a}$$
$$0 = y - x \tag{1.45b}$$

whose LSC is given by (1.41). Thus, different formulations of a DAE can produce different LSC's.

Chapter 2

Standard LSC

2.1 Canonical Forms

We will now begin our analysis of additional dynamics with the linear time invariant DAE

$$Ax' + Bx = f(t) \tag{2.1}$$

where A, B are constant and A is singular. The LSC of (2.1), which is an ODE, will have the general form

$$x' = \Theta x + h(t) \tag{2.2}$$

where Θ is constant. The eigenvalues of Θ are what determines the dynamics of the completion, and thus the behavior of the additional dynamics. Some of these eigenvalues comes from the original DAE and some are created by extra differentiations. In order to identify the additional eigenvalues, it is important to write (2.1) in a way that will expose its original dynamics. Consider the following example,

$$x' = Ax + By + f(t) \tag{2.3a}$$

$$0 = Cx + Dy + g(t) \tag{2.3b}$$

where D is nonsingular so that the DAE is index one. We can write (2.3) as

$$x' = (A - BD^{-1}C)x + f(t) - BD^{-1}g(t) \tag{2.4a}$$
$$y = -D^{-1}Cx - D^{-1}g(t) \tag{2.4b}$$

This is called the state-space form for the DAE (2.3) [23]. We can now see that the dynamics of the DAE are determined by the ODE (2.4a), more specifically, by the eigenvalues of $A - BD^{-1}C$. One completion of (2.4) can now be obtained by differentiating the constraint (2.4b) and substituting (2.4b) for x', which then gives

$$x' = (A - BD^{-1}C)x + f(t) - BD^{-1}g(t) \tag{2.5a}$$
$$y' = -D^{-1}C[(A - BD^{-1}C)x + f(t) - BD^{-1}g(t)] - D^{-1}g'(t) \tag{2.5b}$$

Writing this system in a matrix form we obtain

$$\begin{bmatrix} x \\ y \end{bmatrix}' = \begin{bmatrix} A - BD^{-1}C & 0 \\ -D^{-1}C(A - BD^{-1}C) & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} h_1(t) \\ h_2(t) \end{bmatrix} \tag{2.6}$$

where h_1, h_2 are some functions involving f, g and g'. The eigenvalues of this ODE are given by the eigenvalues of $A - BD^{-1}C$ and 0's. Thus, we conclude that the additional dynamics of (2.5) are given by zero eigenvalues. The implication of this is that the additional dynamics will consist of polynomial expressions.

We will follow a similar process to analyze the additional dynamics of a LSC. The basic idea is to separate the dynamical part of the DAE from the nondynamical part. The dynamical part will remain invariant under the completion process, thus the additional dynamics will come from the differentiation of the nondynamical part. However, we don't have to write the DAE in a semiexplicit form as in the previous example to separate the original dynamics. Consider the following linear DAE

$$\begin{bmatrix} 0 & 2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} x' + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 1 \\ 2 \\ t \end{bmatrix} \tag{2.7}$$

We can write (2.7) as

$$2x_2' + x_3' + x_1 = 1 \tag{2.8a}$$

$$x_3' + x_2 = 2 \tag{2.8b}$$

$$x_3 = t \tag{2.8c}$$

Substituting $x_3 = t$ in (2.8b), we get $x_2 = 2 - 1 = 1$. Then, by substituting x_3 and x_2 in (2.8a), we get $x_1 = 0$. Therefore, $x = x(t) = (0, 1, t)$ is the only solution of (2.8). That is, the solution manifold is zero dimensional and has no dynamics. Thus, (2.8) is equivalent to a pure algebraic system. Note that in this example we have that

$$A = \begin{bmatrix} 0 & 2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

is a nilpotent matrix. In general, for a nilpotent matrix N of index k, the DAE

$$Nx' + x = f$$

has only one solution, which is given by

$$x = (ND + I)^{-1} f = \sum_{i=0}^{k-1} (-1)^i N^i f^{(i)}$$

where $D = d/dt$. Because of this property, nilpotent matrices play a fundamental role in our analysis of linear DAEs due to the canonical decomposition, which we will describe shortly [36], [9], [60].

Suppose that the DAE (2.1) is regular. That is, the matrix pencil $A + \lambda B$ is regular. Then, it has been shown in [36] that there exists nonsingular transformations P and Q such that

$$PAQ = \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}, PBQ = \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix} \tag{2.9}$$

where N is nilpotent of index k. Therefore, pre-multiplication by P and the change of variable $x = Qy$, transforms (2.1) to

$$\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} y' + \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix} y = Pf \tag{2.10}$$

which can be written as

$$y_1' + C_1 y_1 = f_1 \tag{2.11a}$$

$$N y_2' + y_2 = f_2 \tag{2.11b}$$

(2.11a) is an ODE, the general solution of which is given by

$$y_1(t) = e^{-Ct} y_0 + \int_0^t e^{-C(t-s)} f_1(s) ds$$

and the unique solution of (2.11b) is given by

$$y_2(t) = (N\frac{d}{dt} + I)^{-1} f_2 = \sum_{i=0}^{k-1} (-1)^i N^i f_2^{(i)}(t)$$

(2.11) is the form we are going to use in the analysis. Note that the solutions of (2.1) can be obtained from those of (2.11) simply by the transformation $x = Qy$. However, we don't know yet how their LSC's are related. Before we can proceed, we need to establish a relationship between the LSC of the original DAE and that of its canonical form so that (2.11) can be used to analyze the LSC of the original system. Therefore, we will now compare the LSC of

$$Ax' + Bx = f \tag{2.12}$$

with the the LSC of

$$PAQx' + PBQx = Pf \tag{2.13}$$

We will first consider

$$PAx' + PBx = Pf \tag{2.14}$$

where P is invertible to start with. We might need to impose additional restrictions on P to obtain a relationship between the LSC of (2.14) and (2.12). Suppose that (2.12) is solvable with index k. Then, since P is invertible, (2.14) is also solvable with the same index. Now differentiating (2.12) k times with respect to t we get the derivative array

$$J\begin{bmatrix} x' \\ w \end{bmatrix} = -\mathcal{F}x + g \tag{2.15}$$

where

$$J = \begin{bmatrix} A & 0 & 0 & \cdots & 0 \\ B & A & 0 & \cdots & 0 \\ 0 & B & A & \cdots & 0 \\ \vdots & & & \ddots & \vdots \end{bmatrix}, \mathcal{F} = \begin{bmatrix} B \\ 0 \\ 0 \\ \vdots \end{bmatrix}, g = \begin{bmatrix} f \\ f' \\ f'' \\ \vdots \end{bmatrix}$$

Therefore, the LSC of (2.15) is given by the first component solution of the least squares equation (LSE)

$$J^T J\begin{bmatrix} x' \\ w \end{bmatrix} = J^T(-\mathcal{F}x + g) \tag{2.16}$$

Similarly, differentiating (2.14) k times with respect to t we obtain the derivative array

$$\widehat{P}J\begin{bmatrix} x' \\ w \end{bmatrix} = -\widehat{P}\mathcal{F}x + \widehat{P}g \tag{2.17}$$

where

$$\widehat{P} = \begin{bmatrix} P & 0 & 0 & \cdots & 0 \\ 0 & P & 0 & \cdots & 0 \\ 0 & 0 & P & \cdots & 0 \\ \vdots & & & \ddots & \vdots \end{bmatrix}$$

Thus, the least squares equation of (2.17) is

$$J^T\widehat{P}^T\widehat{P}J\begin{bmatrix} x' \\ w \end{bmatrix} = J^T\widehat{P}^T\widehat{P}(-\mathcal{F}x + g) \tag{2.18}$$

We are comparing (2.16) with (2.18). Suppose that P is orthogonal. Then $\widehat{P}$ is orthogonal and $(\widehat{P})^T\widehat{P} = I$, and (2.18) reduces to (2.16). Therefore we can state the following.

Theorem 6 *Suppose that the LSC of (2.12) is given by*

$$x' = \Theta x + h(t) \tag{2.19}$$

Then, the LSC of (2.13), with P orthogonal and Q invertible, is given by

$$y' = Q^{-1}\Theta Q x + \tilde{h}(t) \tag{2.20}$$

Proof. We have already shown that an orthogonal P does not effect the LSC. To see the effect of change of variable, let $x = Qy$. Then (2.19) becomes

$$(Qy)' = \Theta Q y + h$$

so that

$$Qy' = \Theta Q y + h$$

and finally

$$\begin{aligned} y' &= Q^{-1}\Theta Q y + Q^{-1} h \\ &= Q^{-1}\Theta Q y + \tilde{h} \end{aligned} \tag{2.21}$$

□

We are now allowed to use an orthogonal P and any nonsingular Q. Since the canonical form (2.10) is based on nonsingular P, we can not use that form. We have to determine a similar form that can be obtained by orthogonal P and nonsingular Q. First, by (2.10) we have

$$PAQ = \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}, PBQ = \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix}$$

where N is nilpotent of index k, and P, Q are nonsingular. Note that N can be taken strictly upper triangular. Then, by Gram-Schimidt orthogonalization process, there exist a nonsingular upper triangular matrix K such that KP is orthogonal. Thus multiplying the equations by K we get

$$KPAQ = K\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}, KPBQ = K\begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix}$$

Since K is upper triangular we have

$$K\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} = \begin{bmatrix} C_1 & C_2 \\ 0 & N_1 \end{bmatrix}, \quad K\begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} D_1 & D_2 \\ 0 & D_3 \end{bmatrix} \tag{2.22}$$

where C_1, D_1, D_3 are invertible. D_3 is upper triangular and N_1 is strictly upper triangular and thus nilpotent. Let $\tilde{P} = KP$ and $\tilde{Q} = Q$. Then $\tilde{P}$ is orthogonal and $\tilde{Q}$ is invertible, and

$$\tilde{P}A\tilde{Q} = \begin{bmatrix} C_1 & C_2 \\ 0 & N_1 \end{bmatrix}, \tilde{P}B\tilde{Q} = \begin{bmatrix} D_1 & D_2 \\ 0 & D_3 \end{bmatrix} \tag{2.23}$$

Therefore, left multiplication by the orthogonal matrix $\tilde{P}$ and the coordinate change given by $x = \tilde{Q}y$ transforms (2.1) to

$$C_1 y_1' = -C_2 y_2' + D_1 y_1 + D_2 y_2 + f_{11} \tag{2.24a}$$

$$N_1 y_2' = D_3 y_2 + f_{12} \tag{2.24b}$$

Then, using the transformation $y_2 = D_3^{-1} z_2$ and relabeling the coefficients, we obtain the system

$$C_1 y_1' = -C_2 y_2' + D_1 y_1 + D_2 y_2 + f_1 \tag{2.25a}$$

$$N_1 y_2' = y_2 + f_2 \tag{2.25b}$$

where C_1 is nonsingular and N is nilpotent and, in fact, strictly upper triangular. This is the form we are allowed to use. We will now calculate the LSC of (2.25) , analyze its additional dynamics and relate it back to the LSC of the original system (2.1).

Note that (2.25) now has the form

$$F_1(x_1', x_1, x_2', x_2, t) = 0 \tag{2.26a}$$

$$F_2(x_2', x_2, t) = 0 \tag{2.26b}$$

Since C_1 is nonsingular, F_1 is index zero in the variable x_1. The next theorem tells us that the LSC of such a system can be calculated separately, with the first equation being invariant under the LSC process.

Theorem 7 *Suppose that (2.26) is a solvable DAE which satisfies (A1)–(A4). Suppose also that (2.26a) is index zero in* x_1*, that is,* $\partial F_1/\partial x_1'$ *is nonsingular. Then the least squares completion of (2.26) consists of (2.26a) and the least squares completion of (2.26b).*

Proof. Compute the derivative array equations except first list all the derivatives of (2.26a) and then list all the derivatives of (2.26b). Similarily when taking the Jacobians we first partial with respect to x_1' and its higher derivatives w_1 and then with respect to x_2' and its higher derivatives w_2. This modification consists of permutations of the usual least squares equations and thus the new equations are equivalent to the old ones. The least squares equations are then of the form

$$\left[\begin{array}{cc|cc} G_{x_1'} & G_{w_1} & G_{x_2'} & G_{w_2} \\ \hline 0 & 0 & \widetilde{G}_{x_2'} & \widetilde{G}_{w_2} \end{array}\right]^T \left[\begin{array}{c} G(x_1', w_1, x_2', w_2, x_1, x_2, t) \\ \widetilde{G}(x_2', w_2, x_2, t) \end{array}\right] \tag{2.27}$$

$$= \left[\begin{array}{c|c} \Phi_1 & 0 \\ \hline \Phi_2 & [\widetilde{G}_{x_2'}\, \widetilde{G}_{w_2}]^T \end{array}\right] \left[\begin{array}{c} G(x_1', w_1, x_2', w_2, x_1, x_2, t) \\ \widetilde{G}(x_2', w_2, x_2, t) \end{array}\right] = 0 \tag{2.28}$$

where Φ_1 is invertible. Thus we can perform row operations to make $\Phi_1 = I$ and to zero out Φ_2 without changing the solution of the least squares equations. Thus (2.28) is equivalent to

$$G(x_1', w_1, x_2', w_2, x_1, x_2, t) = 0 \tag{2.29}$$

$$\left[\widetilde{G}_{x_2'}\, \widetilde{G}_{w_2}\right]^T \widetilde{G}(x_2', w_2, x_2, t) = 0 \tag{2.30}$$

But (2.26a) is the first block equation in (2.29) and (2.30) are the least squares equations for (2.26b) and the theorem follows. □

2.2 Derivative Array and the Completion

Theorem 7 tells us that the additional dynamics of the LSC are created by the nilpotent system

$$Ny' = y + f \tag{2.31}$$

We will now calculate the LSC of (2.31). Suppose that N has a nilpotency of index k so that $N^{k-1} \neq 0$ but $N^k = 0$. Then, the k step derivative array of (2.31) becomes

$$Jw = -\mathcal{F}x + g \tag{2.32}$$

where

$$J = \begin{bmatrix} N & 0 & 0 & \cdots & 0 \\ I & N & 0 & \cdots & 0 \\ 0 & I & N & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & N \end{bmatrix}, \mathcal{F} = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, g = \begin{bmatrix} f \\ f' \\ f'' \\ \vdots \\ f^{(k)} \end{bmatrix}$$

Then, the least squares equations are given by

$$\begin{bmatrix} N^T & I & 0 & \cdots & 0 & 0 \\ 0 & N^T & I & \cdots & 0 & 0 \\ 0 & 0 & N^T & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & N^T & I \\ 0 & 0 & 0 & \cdots & 0 & N \end{bmatrix} \begin{bmatrix} N & 0 & 0 & \cdots & 0 & 0 \\ I & N & 0 & \cdots & 0 & 0 \\ 0 & I & N & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & N & 0 \\ 0 & 0 & 0 & \cdots & I & N \end{bmatrix} \begin{bmatrix} y' \\ y'' \\ y''' \\ \vdots \\ y^{(k)} \\ y^{(k+1)} \end{bmatrix} \tag{2.33}$$

$$= \begin{bmatrix} N^T & I & 0 & \cdots & 0 & 0 \\ 0 & N^T & I & \cdots & 0 & 0 \\ 0 & 0 & N^T & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & N^T & I \\ 0 & 0 & 0 & \cdots & 0 & N \end{bmatrix} \left\{ - \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} y + \begin{bmatrix} f \\ f' \\ f'' \\ \vdots \\ f^{(k-1)} \\ f^{(k)} \end{bmatrix} \right\} \tag{2.34}$$

Let $M = N^T$. We first multiply both sides by the following nonsingular matrix that corresponds to a series of elementary row operations:

$$K_1 = \begin{bmatrix} I & 0 & 0 & \cdots & 0 & 0 \\ -M & I & 0 & \cdots & 0 & 0 \\ M^2 & -M & I & \cdots & 0 & 0 \\ \vdots & & & & & \\ (-1)^{k-1}M^{k-1} & (-1)^{k-2}M^{k-2} & (-1)^{k-3}M^{k-3} & \cdots & I & 0 \\ 0 & (-1)^{k-1}M^{k-1} & (-1)^{k-2}M^{k-2} & \cdots & -M & I \end{bmatrix}$$

Then (2.34) becomes

$$\begin{bmatrix} M & I & 0 & \cdots & 0 & 0 \\ -M^2 & 0 & I & \cdots & 0 & 0 \\ M^3 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & 0 & I \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} N & 0 & 0 & \cdots & 0 & 0 \\ I & N & 0 & \cdots & 0 & 0 \\ 0 & I & N & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & N & 0 \\ 0 & 0 & 0 & \cdots & I & N \end{bmatrix} \begin{bmatrix} y' \\ y'' \\ y''' \\ \vdots \\ y^{(k)} \\ y^{(k+1)} \end{bmatrix}$$

$$= - \begin{bmatrix} M & I & 0 & \cdots & 0 & 0 \\ -M^2 & 0 & I & \cdots & 0 & 0 \\ M^3 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & 0 & I \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \left\{ - \begin{bmatrix} y \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} y + \begin{bmatrix} f \\ f' \\ f'' \\ \vdots \\ f^{(k-1)} \\ f^{(k)} \end{bmatrix} \right\}$$

which simplifies to

$$\begin{bmatrix} MN+I & N & 0 & \cdots & 0 & 0 \\ -M^2N & I & N & \cdots & 0 & 0 \\ M^3N & 0 & I & \cdots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & I & N \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} y' \\ y'' \\ y''' \\ \vdots \\ y^{(k-1)} \\ y^{(k)} \end{bmatrix} = - \begin{bmatrix} M & I & 0 & \cdots & 0 & 0 \\ -M^2 & 0 & I & \cdots & 0 & 0 \\ M^3 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & 0 & I \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} -f+y \\ -f' \\ -f'' \\ \vdots \\ -f^{(k-1)} \\ -f^{(k)} \end{bmatrix} \tag{2.35}$$

Now multiplying both sides by

$$K_2 = \begin{bmatrix} I & -N & N^2 & \cdots & (-1)^{k-1}N^{k-1} & 0 \\ 0 & I & 0 & \cdots & 0 & 0 \\ 0 & 0 & I & \cdots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & I & 0 \\ 0 & 0 & 0 & \cdots & 0 & I \end{bmatrix}$$

we get

$$\begin{bmatrix} I+XN & 0 & 0 & \cdots & 0 & 0 \\ -M^2N & I & N & \cdots & 0 & 0 \\ M^3N & 0 & I & \cdots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & I & N \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} y' \\ y'' \\ y''' \\ \vdots \\ y^{(k-1)} \\ y^{(k)} \end{bmatrix} \tag{2.36}$$

$$= -\begin{bmatrix} X & I & -N & \cdots & (-1)^{k-2}N^{k-2} & (-1)^{k-1}N^{k-1} \\ -M^2 & 0 & I & \cdots & 0 & 0 \\ M^3 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & 0 & I \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} -f+y \\ -f' \\ -f'' \\ \vdots \\ -f^{(k-1)} \\ -f^{(k)} \end{bmatrix} \tag{2.37}$$

where $X = M + NM^2 + N^2M^3 + \cdots + N^{k-2}M^{k-1}$. Then, since $I + XN$ is invertible, the first block row of (2.37) gives us the ODE

$$(I + XN)y' = -Xy + Xf - \sum_{i=1}^{k}(-1)^i N^{i-1} f^{(i)} \tag{2.38}$$

which is

$$\begin{aligned} y' &= -(I+XN)^{-1}Xy + (I+XN)^{-1}[Xf - \sum_{i=1}^{k}(-1)^i N^{i-1} f^{(i)}] \\ &= -(I+XN)^{-1}Xy + h(t) \end{aligned} \tag{2.39}$$

(2.39) is the LSC of (2.31). Since $h(t)$ is independent of x, the dynamics of (2.39) are determined by the eigenvalues of the matrix $-\Theta$ where

$$\Theta = (I + XN)^{-1}X \tag{2.40}$$

We will now examine the eigenvalues of Θ.

Lemma 1 *Suppose that A is a nilpotent matrix of index k, $B = A^T$ and $X = B + AB^2 + ... + A^{k-2}B^{k-1}$. Then*

$$X(I + AX)^{-1} = (I + XA)^{-1}X \tag{2.41}$$

and $\Theta = (I + XA)^{-1}X$ *is nilpotent.*

Proof. We can write $X + XAX = X + XAX$ so that

$$X(I + AX) = (I + XA)X$$

Multiplying this on the left by $(I + XA)^{-1}$ and on the right by $(I + AX)^{-1}$ we obtain

$$(I + XA)^{-1}X(I + AX)(I + AX)^{-1} = (I + XA)^{-1}(I + XA)X(I + AX)^{-1}$$

which gives us

$$(I + XA)^{-1}X = X(I + AX)^{-1}$$

Thus, we have $\Theta = X(I+AX)^{-1}$. Now let $S = I + NM + N^2M^2 + ... + N^{k-1}M^{k-1}$. Noting that $N^k = M^k = 0$, we have

$$X = M + NM^2 + ... + N^{k-2}M^{k-1} = SM$$

and

$$I + NX = I + N(M + NM^2 + ... + N^{k-2}M^{k-1}) = S$$

which implies

$$\Theta = SMS^{-1}$$

Thus, by the similarity to $M = N^T$, Θ is nilpotent. □

2.3 The Additional Dynamics

We will now analyze what Lemma 1 implies in terms of our original DAE (2.1). Combining (2.39) with Theorem 7, we have proved that the LSC of (2.25) is given by

$$C_1y_1' = -C_2y_2' + D_1y_1 + D_2y_2 + f_1 \tag{2.42a}$$

$$y_2' = -X(I + AX)^{-1}y_2 + h_2(t) \tag{2.42b}$$

We will first show that the solutions of (2.42) move away from the solution manifold of (2.25). Then, we are going to apply the same argument to the original DAE and its LSC.

Now, the solution manifold of (2.25) can be expressed as

$$M = \{(c, x_2(t)) | c \in R^d, Nx_2'(t) + x_2(t) = f_2\}$$

where $d = n - \dim(N)$. Suppose that $y = y(t)$ is a solution of (2.42) and let $y = [y_1^T, y_2^T]^T$. Since (2.42b) is a completion of (2.25b), $x_2 = x_2(t)$ is a particular solution of (2.42b). Therefore, $y_2(t)$ and $x_2(t)$ are two solutions of (2.42b). Thus, by substraction, we have

$$(y_2 - x_2)' = -X(I + AX)^{-1}(y_2 - x_2)$$

That is, $y_2 - x_2$ is a particular solution of the ODE

$$z' = -X(I + AX)^{-1}z$$

Now suppose that $(0, y_2(0))$ is not a consistent initial value. Then, $y_2(0) - x_2(0) \neq 0$ since the set $\{(t, x_2(t))\}$ defines the solution manifold. Note that any solution of the ODE $z' = \Theta z$ where Θ is nilpotent, is of the form

$$z(t) = (I + \Theta t + \frac{1}{2}\Theta^2 t^2 + \cdots + \frac{1}{k!}\Theta^{k-1}t^{k-1})c$$

Therefore since $\Theta = X(I + AX)^{-1}$ is nilpotent, $y_2(t) - x_2(t)$ never goes to zero if $y_2(0) - x_2(0) \neq 0$. Thus, if $\Theta(y_2(0) - x_2(0) \neq 0$, then

$$\|y_2(t) - x_2(t)\| \to \infty \tag{2.43}$$

and

$$d(y(t), M) \to \infty \tag{2.44}$$

In other words, the solution of (2.42) starting near the solution manifold (but not on it) will gradually move further away from the manifold, although at polynomial speed.

Lets now see what this means in terms of the original system. Suppose that we start with a solvable DAE

$$Ax' + Bx = f \tag{2.45}$$

Let

$$y' = \Theta y + h(t) \tag{2.46}$$

be the LSC of (2.45). Let M be the solution manifold of (2.45). Then, the set $Q^{-1}M = \{Q^{-1}x | x \in M\}$ will be the solution manifold of (2.25). Now let $y = y(t)$ be a solution of (2.46). Then, by Theorem 6, $Q^{-1}y$ will be a solution of (2.42). Then, by the foregoing argument, we have

$$d(Q^{-1}y(t), Q^{-1}M) \to \infty$$

provided that $(Q^{-1}y)(0)$ is not consistent for (2.25) and $\Theta Q^{-1}y(0) \neq 0$. Since Q is a constant matrix, thus bounded, this implies that

$$d(y(t), M) \to \infty$$

as $t \to \infty$ provided that $y(0)$ is not consistent for (2.45) and $\Theta Q^{-1}y(0) \neq 0$.

Chapter 3

Stabilized LSC

3.1 Stabilized Differentiation

In the previous section we have determined the additional dynamics of the LSC defined by the standard derivative array, one that is formed by successively differentiating the DAE. We have concluded that the solution manifold is not stable in that case if the index is greater than 1. One way to change this outcome is to modify the derivative array equations in some fashion before solving them in the least squares sense. A technique known as Baumgarte stabilization [7] has been used to stabilize the constraints of ODE's. It is based on connecting constraints using a parameter during the differentiation. The parameter is later assigned an appropriate value to achieve the desired stability. There can be technical issues in the selection process of the parameter [2], however, Baumgarte stabilization is often used in practice.

To illustrate the idea, consider the following pure algebraic DAE

$$Ax + f(t) = 0 \tag{3.1}$$

where A is nonsingular. Suppose that we want to embed (3.1) in an ODE as an asymptotically stable set. Let λ be a complex parameter. Now consider the following equation

$$(Ax + f(t))' + \lambda((Ax + f(t)) = 0 \tag{3.2}$$

This is an ODE and it contains the solutions of (3.1). Let $z = z(t)$ be a solution of (3.2).

Then, $u = Az + f$ satisfies the ODE

$$u' + \lambda I u = 0 \tag{3.3}$$

Therefore we have $u = e^{-\lambda I t} u_0$. This implies that $u \to 0$ for any real $\lambda > 0$. But since $u = Az + f$, we get $Az + f \to 0$ and thus $z(t) \to x(t)$ since A is constant, where $x(t)$ is the solution of (3.1). In other words, the solutions of the completion (3.2) converge to the solution manifold of the underlying DAE (3.1).

We want to generalize this idea for DAEs having a general structure and arbitrary index. Consider the following system

$$\begin{aligned} F &= 0 && \text{(3.4a)} \\ (\frac{d}{dt} + \lambda)F &= 0 && \text{(3.4b)} \\ &\vdots && \text{(3.4c)} \\ (\frac{d}{dt} + \lambda)^k F &= 0 && \text{(3.4d)} \end{aligned}$$

where $F = Ax' + Bx - f$. Instead of just differentiating the DAE, we add a λ multiple of all the previous equations. This is called stabilized differentiation. Let

$$D = \begin{bmatrix} I & 0 & 0 & \cdots & 0 \\ \lambda & I & 0 & \cdots & 0 \\ \lambda^2 & 2\lambda & I & \cdots & 0 \\ \vdots & & & & \\ \lambda^k & k\lambda^{k-1} & (k(k-1)/2)\lambda^{k-2} & \cdots & I \end{bmatrix}. \tag{3.5}$$

Then, the equations (3.4) can be expressed as

$$DJ \begin{bmatrix} x' \\ w \end{bmatrix} = D(-\mathcal{F}y + g). \tag{3.6}$$

where $J, \mathcal{F}, g$ are the same quantities corresponding to the standard derivative array defined in the previous chapter. We have calculated the previous LSC using elementary matrix operations. Because of the additional complexity created by the presence of D

in the equation, we will employ a different technique in this section. Note that because of the uniqueness of the LSC, the LSC of (3.4) can be also obtained as the first block row of $(DJ)^{\dagger}D(-\mathcal{F}y+g)$, where † denotes the Moore-Penrose inverse [24]. That is,

$$x' = V^T(DJ)^{\dagger}D(-\mathcal{F}y+g) \tag{3.7}$$

where $V^T = \begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix}$.

Theorem 8 *Suppose that we have a constant coefficient DAE*

$$Ay' + By = f \tag{3.8}$$

that is solvable of index k, with the derivative array

$$J \begin{bmatrix} x' \\ w \end{bmatrix} = -\mathcal{F}y + g. \tag{3.9}$$

Suppose that (A1)–(A4) are satisfied for this system. Let G_0 be an $n \times (k+1)n$ matrix satisfying

$$G_0 J = \begin{bmatrix} I & 0 & 0 \cdots 0 \end{bmatrix} \tag{3.10a}$$

$$G_0 Z = 0 \tag{3.10b}$$

where Z is a matrix of maximal rank satisfying $Z^T J = 0$. Namely, the columns of Z form a basis for $N(J^T)$. Then, the least squares completion of (3.8) defined by (3.9) is

$$y' = G_0(-\mathcal{F}y+g) \tag{3.11}$$

Proof. Let

$$G_0 = \begin{bmatrix} X_0 & X_1 & X_2 \cdots X_k \end{bmatrix}$$

and

$$Z^T = \begin{bmatrix} Z_0^T & Z_1^T & Z_2^T \cdots Z_k^T \end{bmatrix}$$

Suppose that rank$(Z) = a$. Then corank$(J) = a$. By the Gram-Schmidt method, there exists nonsingular matrices D_1, D_2 such that the matrices $D_1 G$ and $D_2 Z$ have orthonormal set of row vectors separately, where

$$D_1 G = \begin{bmatrix} D_1 X_0 & D_1 X_1 & D_1 X_2 \cdots D_1 X_k \end{bmatrix}$$

and

$$D_2 Z^T = \begin{bmatrix} D_2 Z_0^T & D_2 Z_1^T & D_2 Z_2^T \cdots D_2 Z_k^T \end{bmatrix}$$

Then, $GZ = 0$ implies

$$(D_1 G)(D_2 Z^T)^T = \begin{bmatrix} D_1 X_0 & D_1 X_1 & D_1 X_2 \cdots D_1 X_k \end{bmatrix} \begin{bmatrix} D_2 Z_0^T & D_2 Z_1^T & D_2 Z_2^T \cdots D_2 Z_k^T \end{bmatrix}^T$$

$$= \begin{bmatrix} D_1 X_0 & D_1 X_1 & D_1 X_2 \cdots D_1 X_k \end{bmatrix} \begin{bmatrix} Z_0 D_2^T \\ Z_1 D_2^T \\ Z_2 D_2^T \\ \vdots \\ Z_k D_2^T \end{bmatrix}$$

$$= D_1 (GZ) D_2^T = 0$$

Thus, the matrix

$$\begin{bmatrix} D_1 G \\ D_2 Z^T \end{bmatrix}$$

has an orthonormal set of row vectors. Therefore, we can find a matrix R such that

$$Q = \begin{bmatrix} D_1 G \\ D_2 Z^T \\ R \end{bmatrix}$$

is an orthogonal matrix. Note that the least squares completion of (3.8) is given by

$$y' = J^\dagger(-\mathcal{F}y + g)$$

On the other hand, since $Q^T Q = I$, we have from [24] that

$$(QJ)^\dagger(-Q\mathcal{F} + Qg)$$

$$= J^\dagger Q^T(-Q\mathcal{F} + Qg)$$

$$J^\dagger(-\mathcal{F} + g)$$

That is, multiplying the system by Q does not change the least squares solution. Now, in a more compact form, we can write $[QJ, Q(-\mathcal{F}y+g)]$ as

$$\left[\begin{array}{c|cccccc|c} D_1 & 0 & 0 & \cdots & 0 & 0 & & D_1G(-\mathcal{F}+g) \\ 0 & 0 & 0 & \cdots & 0 & 0 & & D_2Z^T(-\mathcal{F}+g) \\ \hline M_0 & M_1 & M_2 & \cdots & & & & \cdots \end{array}\right] \tag{3.12}$$

Since corank(QJ) = corank(J) = a = rank(Z), which is the size of the zero block row, we can choose R so that $M_1, M_2, ..., M_k$ block row has full row rank. In this special circumstance we have from [24] that

$$\begin{bmatrix} X & 0 \\ Y & Z \end{bmatrix}^{\dagger} = \begin{bmatrix} X^{\dagger} & 0 \\ -X^{\dagger}YZ^{\dagger} & Z^{\dagger} \end{bmatrix}$$

Then, the first block row of (3.12) produces the least squares completion

$$\begin{aligned} y' &= D_1^{-1}\left[D_1G_0(-\mathcal{F}y+g)\right] \\ &= G_0(-\mathcal{F}y+g) \end{aligned} \tag{3.13}$$

□

3.2 Derivative Array and the Completion

We will now investigate the additional dynamics of the LSC defined by (3.4) using this lemma. An easy calculation shows that Theorem 6 and Theorem 7 hold the same way for the derivative array (3.4). Thus, we only need to consider the nilpotent system

$$Nx' + x = f \tag{3.14}$$

Suppose that N has an index of k. Then, differentiating (3.14) k times with respect to t in the sense of (3.4), we get the derivative array

$$DJ\begin{bmatrix} x' \\ w \end{bmatrix} = D(-\mathcal{F}x+g) \tag{3.15}$$

where

$$J = \begin{bmatrix} N & 0 & 0 & \cdots & 0 \\ I & N & 0 & \cdots & 0 \\ 0 & I & N & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & N \end{bmatrix}, \mathcal{F} = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, g = \begin{bmatrix} f \\ f' \\ f'' \\ \vdots \\ f^{(k)} \end{bmatrix}$$

We will apply Theorem 8 to calculate the actual completion. Therefore we need to find a matrix G_0 that satisfies

$$G_0(DJ) = \begin{bmatrix} I & 0 & 0 \cdots 0 \end{bmatrix} \tag{3.16a}$$

$$G_0 Z = 0 \tag{3.16b}$$

where Z is a matrix of maximal rank with $Z^T(DJ) = 0$.

Lets first write

$$J = \begin{bmatrix} N & 0 & 0 & \cdots & 0 \\ I & N & 0 & \cdots & 0 \\ 0 & I & N & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & N \end{bmatrix} = \begin{bmatrix} N & 0 & 0 & \cdots & 0 \\ 0 & N & 0 & \cdots & 0 \\ 0 & 0 & N & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & N \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ I & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{3.17}$$

$$= J_N + J_I \tag{3.18}$$

Since J_N is block diagonal with the same diagonal block entries and the elements of D

are the scalar multiples of I, we have that $DJ_N = J_N D$. Therefore,

$$\begin{aligned}
DJD^{-1} &= D\begin{bmatrix} N & 0 & 0 & \cdots & 0 \\ 0 & N & 0 & \cdots & 0 \\ 0 & 0 & N & \cdots & 0 \\ \vdots \\ 0 & 0 & 0 & \cdots & N \end{bmatrix} D^{-1} + D\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ I & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & \cdots & 0 \\ \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} D^{-1} && (3.19) \\
&= \begin{bmatrix} N & 0 & 0 & \cdots & 0 \\ 0 & N & 0 & \cdots & 0 \\ 0 & 0 & N & \cdots & 0 \\ \vdots \\ 0 & 0 & 0 & \cdots & N \end{bmatrix} + D\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ I & 0 & 0 & \cdots & 0 \\ 0 & I & 0 & \cdots & 0 \\ \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} D^{-1} && (3.20) \\
&= \begin{bmatrix} N & 0 & 0 & \cdots & 0 \\ 0 & N & 0 & \cdots & 0 \\ 0 & 0 & N & \cdots & 0 \\ \vdots \\ 0 & 0 & 0 & \cdots & N \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ I & 0 & 0 & \cdots & 0 \\ \lambda I & I & 0 & \cdots & 0 \\ \vdots \\ \lambda^{k-1} I & \lambda^{k-2} I & \lambda^{k-3} I & \cdots & 0 \end{bmatrix} && (3.21) \\
&= \begin{bmatrix} N & 0 & 0 & \cdots & 0 \\ I & N & 0 & \cdots & 0 \\ \lambda I & I & N & \cdots & 0 \\ \vdots \\ \lambda^{k-1} I & \lambda^{k-2} I & \lambda^{k-3} I & \cdots & N \end{bmatrix}. && (3.22)
\end{aligned}$$

Note that we have

$$D^{-1} = \begin{bmatrix} I & 0 & 0 & \cdots & 0 \\ -\lambda & I & 0 & \cdots & 0 \\ \lambda^2 & -2\lambda & I & \cdots & 0 \\ \vdots \\ \lambda^k & k\lambda^{k-1} & (k(k-1)/2)\lambda^{k-2} & \cdots & I \end{bmatrix}.$$

Let $G_0 = \begin{bmatrix} X_0 & X_1 & X_2 \cdots X_k \end{bmatrix}$. Then, multiplying both sides of (3.22) by G_0 and using (3.16a) we get

$$G_0 DJD^{-1} = G_0 \begin{bmatrix} N & 0 & 0 & \cdots & 0 \\ I & N & 0 & \cdots & 0 \\ \lambda I & I & N & \cdots & 0 \\ \vdots & & & & \\ \lambda^{k-1} I & \lambda^{k-2} I & \lambda^{k-3} I & \cdots & N \end{bmatrix}$$

so that

$$\begin{bmatrix} I & 0 & 0 \cdots 0 \end{bmatrix} D^{-1} = G_0 \begin{bmatrix} N & 0 & 0 & \cdots & 0 \\ I & N & 0 & \cdots & 0 \\ \lambda I & I & N & \cdots & 0 \\ \vdots & & & & \\ \lambda^{k-1} I & \lambda^{k-2} I & \lambda^{k-3} I & \cdots & N \end{bmatrix}$$

and hence

$$\begin{bmatrix} I & 0 & 0 \cdots 0 \end{bmatrix} = G_0 \begin{bmatrix} N & 0 & 0 & \cdots & 0 \\ I & N & 0 & \cdots & 0 \\ \lambda I & I & N & \cdots & 0 \\ \vdots & & & & \\ \lambda^{k-1} I & \lambda^{k-2} I & \lambda^{k-3} I & \cdots & N \end{bmatrix} \tag{3.23}$$

which produces the following equations

$$\begin{aligned} X_0 N + X_1 + \lambda X_2 + \lambda^2 X_3 + \lambda^3 X_4 \cdots \lambda^{k-1} X_k &= I && (3.24a) \\ X_1 N + X_2 + \lambda X_3 + \lambda^2 X_4 + \cdots \lambda^{k-2} X_k &= 0 && (3.24b) \\ &\vdots && (3.24c) \\ X_{k-3} N + X_{k-2} + \lambda X_{k-1} + \lambda^2 X_k &= 0 && (3.24d) \\ X_{k-2} N + X_{k-1} + \lambda X_k &= 0 && (3.24e) \\ X_{k-1} N + X_k &= 0 && (3.24f) \\ X_k N &= 0 && (3.24g) \end{aligned}$$

Multiplying each equation by $-\lambda$ and adding it to the equation directly above it, we get

$$\begin{aligned}
X_k &= -X_{k-1}N(I-\lambda N)^{-1} && (3.25a)\\
X_{k-1} &= -X_{k-2}N(I-\lambda N)^{-1} && (3.25b)\\
X_{k-2} &= -X_{k-3}N(I-\lambda N)^{-1} && (3.25c)\\
&\vdots && (3.25d)\\
X_3 &= -X_2N(I-\lambda N)^{-1} && (3.25e)\\
X_2 &= -X_1N(I-\lambda N)^{-1} && (3.25f)\\
X_1 &= (I-X_0N)(I-\lambda N)^{-1} && (3.25g)
\end{aligned}$$

Let $U=(I-\lambda N)^{-1}$ and $V=N(I-\lambda N)^{-1}$. Then

$$\begin{aligned}
G_0 &= \begin{bmatrix} X_0 & X_1 & X_2\cdots X_k \end{bmatrix} && (3.26)\\
&= \begin{bmatrix} X_0 & U-X_0V & -(U-X_0V)V & (U-X_0V)V^2\cdots \end{bmatrix} && (3.27)\\
&= X_0\begin{bmatrix} I & -V & V^2 & -V^3\cdots & (-1)^{k-1}V^{k-1} \end{bmatrix} && (3.28)\\
&+ U\begin{bmatrix} 0 & I & -V & V^2\cdots & (-1)^{k-2}V^{k-2} \end{bmatrix} && (3.29)
\end{aligned}$$

Note that V is also nilpotent of index k since $(I-\lambda N)^{-1}=I+\lambda N+\lambda^2N^2+...+\lambda^{k-1}N^{k-1}$.

We will use (3.16b) to determine X_0. Let Z be a matrix of maximal rank satisfying $Z^T(DJ)=0$. Then, multiplying (3.22) by Z^T gives

$$Z^TDJD^{-1}=Z^T\begin{bmatrix} N & 0 & 0 & \cdots & 0\\ I & N & 0 & \cdots & 0\\ \lambda I & I & N & \cdots & 0\\ \vdots & & & & \\ \lambda^{k-1}I & \lambda^{k-2}I & \lambda^{k-3}I & \cdots & N \end{bmatrix} \tag{3.30}$$

Let $Z^T=\begin{bmatrix} Z_0^T & Z_1^T & Z_2^T & \cdots Z_k^T \end{bmatrix}$. Then by a similar argument, (3.30) implies

$$\begin{aligned}
Z_k^T &= -Z_{k-1}^T N(I-\lambda N)^{-1} && (3.31a)\\
Z_{k-1}^T &= -Z_{k-2}^T N(I-\lambda N)^{-1} && (3.31b)\\
&\vdots && (3.31c)\\
Z_2^T &= -Z_1 N^T(I-\lambda N)^{-1} && (3.31d)\\
Z_1^T &= -Z_0 N^T(I-\lambda N)^{-1} && (3.31e)
\end{aligned}$$

Then for $Z_0 = I$, Z will have maximal rank and

$$Z = \begin{bmatrix} I & -V & V^2 & \cdots V^{k-1} \end{bmatrix}^T.$$

where $V = N(I-\lambda N)^{-1}$. Therefore (3.16b) implies that $GZ = 0$ gives

$$\left(X_0 \begin{bmatrix} I & -V & V^2 & -V^3\cdots \end{bmatrix} + U \begin{bmatrix} 0 & I & -V & V^2\cdots \end{bmatrix}\right) \left(\begin{bmatrix} I \\ -W \\ W^2 \\ -W^3\cdots \end{bmatrix}\right) = 0$$

where $W = V^T$. From this we obtain

$$\begin{aligned}
X_0 &= U(W + VW^2 + \cdots + V^{k-2}W^{k-1})(I + VW + V^2W^2 + \cdots + V^{k-1}W^{k-1})^{-1}\\
&= UX(I+VX)^{-1} && (3.32)
\end{aligned}$$

where $X = W + VW^2 + V^2W^3 + ..V^{k-2}W^{k-1}$. On the other hand we have

$$\lambda \begin{bmatrix} N \\ I \\ \lambda I \\ \vdots \\ \lambda^{k-2}I \\ \lambda^{k-1}I \end{bmatrix} - \begin{bmatrix} I \\ \lambda I \\ \lambda^2 I \\ \vdots \\ \lambda^{k-1}I \\ \lambda^k I \end{bmatrix} = \begin{bmatrix} \lambda N - I \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \qquad (3.33)$$

Multiplying both sides of (3.33) by G_0 and using (3.23), we get

$$\lambda I - G_0\mathcal{F} = X_0(\lambda N - I) \tag{3.34}$$

so that

$$\begin{aligned} G_0\mathcal{F} &= \lambda I + x_0(I - \lambda N) \\ &= \lambda I + UX(I + VX)^{-1}U^{-1} \end{aligned} \tag{3.35}$$

since $U = (I - \lambda N)^{-1}$. Therefore the LSC becomes

$$\begin{aligned} y' &= G_0(-\mathcal{F}y + g) && (3.36) \\ &= [-\lambda I - UX(I + VX)^{-1}U^{-1}]y + h(t) && (3.37) \end{aligned}$$

for some function $h(t)$. The dynamics of this new completion is now determined by

$$\Theta = -\lambda I - UX(I + VX)^{-1}U^{-1} \tag{3.38}$$

Note that $UX(I+VX)^{-1}U^{-1}$ is similar to $X(I+VX)^{-1}$ and $X(I+VX)^{-1}$ is similar to V as proved in the previous section, which is nilpotent. Thus, $UX(I + VX)^{-1}U^{-1}$ is nilpotent. Therefore the eigenvalues of Θ consists of $-\lambda$. This implies, by a similar argument as in the previous chapter, that the additional dynamics will go to zero exponentially for any $\lambda > 0$. In other words, the additional solutions will converge to the solution manifold. Note however, that the additional dynamics are a polynomial times an exponential since $H = UX(I + VX)^{-1}U^{-1}$ is nilpotent. That is, a typical solution will be in the form

$$z(t) = e^{-\lambda t}(I - Ht + \cdots + \frac{(-1)^{k-1}}{(k-1)!}H^{k-1}t^{k-1})c_0$$

For a numerical method, if $k \geq 2$ and if $Re(\lambda)$ is not big enough, this could mean that error could grow before being damped.

Chapter 4

Alternative Stabilized Completion

4.1 Index One Formulation

We have developed a method to obtain a LSC with desired additional eigenvalues and thus desired additional dynamics. However, the dynamics are polynomial times exponentials and not just pure exponentials. We will now introduce an alternative way to obtain a stabilized completion which can give pure exponentials as the additional dynamics. In [42], a method has been developed to obtain an index-1 system from a general DAE. We will combine this idea with stabilized differentiation to produce another completion with desired stability characteristics. Suppose that the linear time invariant DAE

$$Ax' + Bx = f \tag{4.1}$$

is solvable. Then we have

$$PAQ = \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}, PBQ = \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix} \tag{4.2}$$

where P, Q are nonsingular. Let $\tilde{A} = PAQ, \tilde{B} = PBQ, \tilde{f} = Pf$, and $\tilde{J}$ denote the Jacobian of the system

$$\tilde{A}x' + \tilde{B}x = \tilde{f} \tag{4.3}$$

Lemma 2 *corank*$(\tilde{J}) =$ *dim*$(N) = a$.

Proof. Note that the rows of $\tilde{J}$ consists of rows of $\tilde{J}_I$ and $\tilde{J}_N$, where $\tilde{J}_I$ is the Jacobian corresponding to $x' + Cx = h_1$ and $\tilde{J}_N$ is the jacobian corresponding to $Nx' + x = h_2$. $\tilde{J}_I$ obviously has full row rank, therefore it is sufficient to consider the corank of $\tilde{J}_N$. We then have

$$\tilde{J}_N = \begin{bmatrix} N & 0 & 0 & \cdots & 0 \\ I & N & 0 & \cdots & 0 \\ 0 & I & N & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & N \end{bmatrix}_{(k+1)\times(k+1)}$$

We claim that corank$(\tilde{J}_N) =$ dim(N). Suppose that $\tilde{Z}_2$ is a matrix of maximal rank with $\tilde{Z}_2^T J_N = 0$. Let $\tilde{Z}_2^T = \left[\begin{array}{cccc} Z_{2,0}^T & Z_{2,1}^T & Z_{2,2}^T & \cdots Z_{2,k}^T \end{array}\right]$. Then $\tilde{Z}_2^T J_N = 0$ implies

$$0 = Z_{2,0}^T N + Z_{2,1}^T \tag{4.4a}$$
$$0 = Z_{2,1}^T N + Z_{2,2}^T \tag{4.4b}$$
$$\vdots \tag{4.4c}$$
$$0 = Z_{2,k-1}^T N + Z_{2,k}^T \tag{4.4d}$$
$$0 = Z_{2,k}^T N \tag{4.4e}$$

From this we get $\tilde{Z}_2^T = \left[\begin{array}{cccc} Z_{2,0}^T & -Z_{2,0}^T N & Z_{2,0}^T N^2 & \cdots Z_{2,0}^T N^k \end{array}\right]$. Therefore, the choice $Z_{2,0} = I$ renders $\tilde{Z}_2$ maximal with rank$(\tilde{Z}_2) =$ dim(N). Thus we get

$$\text{corank}(\tilde{J}_N) = \text{corank}(\tilde{J}) = \text{dim}(N) = a.$$

□

Now define

$$\widehat{P} = \begin{bmatrix} P & 0 & 0 & \cdots & 0 \\ 0 & P & 0 & \cdots & 0 \\ 0 & 0 & P & \cdots & 0 \\ \vdots & & & \ddots & \vdots \end{bmatrix}$$

Note that $\widehat{P}$ is also invertible. Then an easy calculation shows that

$$\tilde{J} = \widehat{P} J \widehat{Q}$$

Let $Z_2^T = \tilde{Z}_2^T (\widehat{P})^{-1}$. Then, Z_2 has full column rank, $Z_2^T J = 0$ and

$$\text{rank}(Z_2) = \text{rank}(\tilde{Z}_2)$$

Thus we obtain

$$\text{corank}(J) = \text{corank}(\tilde{J}) = \dim(N) = a$$

Now, since $[J, \mathcal{F}]$ has full row rank and $Z_2^T J = 0$, $Z_2^T \mathcal{F}$ has full row rank. Then there exists a T_2 of full column rank such that $Z_2^T \mathcal{F} T_2 = 0$ and $\text{rank}(AT_2) = d = n - a$ [42]. The columns of T_2 set up coordinates for the solution manifold. Let $Z_{1,0}$ be a matrix whose columns form a basis for $R(AT_2)$. Define $Z_1^T = \begin{bmatrix} Z_{1,0}^T & 0 & 0 & \cdots & 0 \end{bmatrix}$.

Lemma 3 $\begin{bmatrix} Z_{1,0}^T A \\ Z_2^T \mathcal{F} \end{bmatrix}$ *is an invertible matrix.*

Proof. Suppose that $\begin{bmatrix} Z_{1,0}^T A \\ Z_2^T \mathcal{F} \end{bmatrix} v = 0$ for some vector v. Then, we have $(Z_{1,0}^T A)v = 0$ and $(Z_2^T \mathcal{F})v = 0$. $(Z_2^T \mathcal{F})v = 0$ implies $v \in N(Z_2^T \mathcal{F}) = R(T_2)$ by definition. Therefore, $v = T_2 x$ for some vector x. Then, by the first part, we get $(Z_{1,0}^T A T_2)x = 0 = Z_{1,0}^T (AT_2 x)$, which implies $(AT_2)x = 0$ since $Z_{1,0}$ is a basis for $R(AT_2)$. Then $x = 0$ since AT_2 has full column rank. Thus, $v = T_2 x = 0$. □

Now consider the following system.

$$Z_{1,0}^T A x' = Z_{1,0}^T(-Bx + f) \tag{4.5a}$$

$$0 = Z_2^T(\mathcal{F}x - g) \tag{4.5b}$$

The equation (4.5b) gives us all the constraints of the original DAE. Then, by the lemma, (4.5) is a semi-explicit index one DAE with the same solutions as (4.1). Thus, any completion of this system will be a completion of the original DAE. We can now differentiate the constraint part in the stabilized sense to get the completion

$$
\begin{aligned}
Z_{1,0}^T A x' &= Z_{1,0}^T(-Bx + f) && \text{(4.6a)} \\
Z_2^T(\mathcal{F}x - g) &= -\lambda Z_2^T(\mathcal{F}x - g) && \text{(4.6b)}
\end{aligned}
$$

More explicitly this is

$$
\begin{bmatrix} Z_{1,0}^T A \\ -Z_2^T \mathcal{F} \end{bmatrix} x' = - \begin{bmatrix} Z_{1,0}^T B \\ \lambda Z_2^T \mathcal{F} \end{bmatrix} x + \begin{bmatrix} Z_{1,0}^T f \\ \lambda Z_2^T g \end{bmatrix} \tag{4.7}
$$

As can be easily seen from the foregoing discussion, the additional eigenvalues of this new completion are also given by $-\lambda$. Moreover, there is no nilpotent part. Therefore, the additional dynamics are now given by pure exponentials. Again, the solution manifold is asymptotically stable for $\lambda > 0$.

For computational reasons, we will now show that the completion (4.6) can be calculated by the least squares method. We will do this by constructing a derivative array and showing that the LSC obtained by the derivative array produces the same completion. We do this to avoid the derivative of Z_2, which appears in the time variable case. Since Z_2 is numerically calculated, such a differentiation could produce too much numerical error.

4.2 Computation using Least Squares

Lets first defines the following shifting operators

$$
S = \begin{bmatrix} 0 & I_{kn} \end{bmatrix}_{(kn)\times(k+1)n}, K = \begin{bmatrix} I_{kn} & 0 \end{bmatrix}_{(kn)\times(k+1)n}, V = \begin{bmatrix} I_n & 0 \end{bmatrix}_{n\times(kn)}
$$

Let $\tilde{J}, \tilde{\mathcal{F}}, \tilde{g}$ denote the same matrices as $J, \mathcal{F}, g$ with $k+1$ differentiations. We then have the following relationships:

$$K\tilde{J} = JK \tag{4.8}$$

$$K\tilde{\mathcal{F}} = \mathcal{F} \tag{4.9}$$

$$K\tilde{g} = g \tag{4.10}$$

$$S\tilde{\mathcal{F}} = 0 \tag{4.11}$$

$$S\tilde{g} = g' \tag{4.12}$$

Let $Z_3^T = \lambda Z_2^T K + Z_2^T S$. Now consider the matrix

$$R = \begin{bmatrix} Z_1^T K \\ Z_2^T K \\ Z_3^T \\ Z_4^T \end{bmatrix}$$

where Z_4^T are extra rows to make R invertible. We will consider the LSC of

$$RJw = R(-\mathcal{F}x + g) \tag{4.13}$$

which is given by the first block row of $(RJ)^\dagger R(-\mathcal{F}x + g)$ where † denotes the Moore-Penrose inverse of the matrix inside [24]. In a compact form we have

$[RJ, R(-\mathcal{F}x + g)]$ is

$$\left[\begin{array}{c|cccccc|c} Z_{1,0}^T A & 0 & 0 & \cdots & 0 & 0 & Z_{1,0}^T(Bx + f) \\ 0 & 0 & 0 & \cdots & 0 & 0 & Z_2^T(\mathcal{F}x + g) \\ -Z_{2,0}^T B & 0 & 0 & \cdots & 0 & 0 & Z_2^T g' + \lambda Z_2^T(\mathcal{F}x + g) \\ \hline 0 & R_1 & R_2 & \cdots & \cdots & \cdots & \cdots \end{array}\right] \tag{4.14}$$

Note that $\begin{bmatrix} Z_{1,0}^T A \\ -Z_{2,0}^T B \end{bmatrix} = \begin{bmatrix} Z_1^T K \\ Z_3^T \end{bmatrix} \tilde{J} V^T$, invertible, and $Z_2^T K \tilde{J} = 0$. Thus the rows of Z_1^T, Z_2^T, Z_3^T are linearly independent. Also, since $\tilde{J}V^T$ has n columns, we get

$$n = \text{rank}(\begin{bmatrix} Z_1^T K \\ Z_2^T K \\ Z_3^T \end{bmatrix} \tilde{J}V^T) \leq \text{rank}(\tilde{J}V^T) \leq n$$

Therefore,

$$\text{corank}(\tilde{J}V^T) = (k+2)n - \text{rank}(\tilde{J}V^T) = (k+1)n = \dim(Z_2K) + \dim(Z_4)$$

This implies that we can choose Z_4 such that both R is invertible and $Z_4^T(\tilde{J}V^T) = R_0 = 0$. In this special circumstance we have from [24] that

$$\begin{bmatrix} X & 0 \\ Y & Z \end{bmatrix}^{\dagger} = \begin{bmatrix} X^{\dagger} & 0 \\ -X^{\dagger}YZ^{\dagger} & Z^{\dagger} \end{bmatrix}$$

so that the Moore-Penrose inverse of the block lower triangular matrix is also block lower triangular matrix. Thus the first block row gives the LSC

$$\begin{bmatrix} Z_{1,0}^T A \\ -Z_{2,0}^T B \end{bmatrix} x' = \begin{bmatrix} Z_{1,0}^T B \\ \lambda Z_2^T \mathcal{F} \end{bmatrix} x + \begin{bmatrix} Z_{1,0}^T f \\ \lambda Z_2^T g \end{bmatrix} \tag{4.15}$$

which is the same as

$$Z_{1,0}^T A x' = Z_{1,0}^T(-Bx + f) \tag{4.16a}$$

$$Z_2^T(-\mathcal{F}x + g)' = (-\lambda I) Z_2^T(-\mathcal{F}x + g) \tag{4.16b}$$

since $Z_2^T \mathcal{F} = Z_{2,0}^T B$.

Chapter 5

Stabilized LSC: LTV Systems

5.1 Introduction

We have developed two methods for constant coefficient DAEs that produce LSC with desired additional dynamics. In this chapter we will apply those techniques to the linear time varying DAE

$$A(t)x' + B(t)x = f(t) \tag{5.1}$$

where $A(t), B(t)$ are matrix functions of t and $A(t)$ is singular possibly for all t. The LTV analysis contains additional difficulties compared to the constant coefficient case. For example, the stability of the system is no longer directly related to the eigenvalues. Ensuring the smoothness of various components during calculations is also a problem since they are done at every time and not just once. Since the derivative array equations will now contain the derivatives of A and B, it is not feasible to obtain an explicit formula of the completion as in the constant coefficient case. Instead, by using the techniques developed in Chapter 3 and 4 together, we will obtain a less explicit but more effective formulation of the completion that will enable us to identify the additional dynamics.

Before starting our analysis, lets take a look at what happens with the standard derivative array in the LTV case. In the LTI case we have proved that the additional eigenvalues are equal to 0. Consider the example

Example 3 *Let α, β be parameters in the index one DAE;*

$$x_1' = \beta x_1 \tag{5.2a}$$

$$0 = e^{\alpha t}(x_1 - x_2) \tag{5.2b}$$

Equation (5.2) is solvable with solutions $x_1 = e^{\beta t}c,\ x_2 = x_1$ for all values of α. Here c is an arbitrary constant. The least squares completion is

$$x_1' = \beta x_1 \tag{5.3a}$$

$$x_2' = -\alpha x_2 + (\alpha + \beta)x_1 \tag{5.3b}$$

The eigenvalues of the system (5.3) are $\{\beta, -\alpha\}$. Note that β comes from the dynamics of the DAE but $-\alpha$ is from the additional dynamics of the LSC.

The example shows that the additional dynamics can be arbitrary in the LTV case if one uses the standard derivative array.

5.2 Derivative Array and Canonical Forms

From here on, we may skip writing the variable t, while all the matrices involved will be assumed time variable unless otherwise specified. Similarly, full rank or invertible, will mean pointwise full rank or pointwise invertible when time variable matrices are concerned.

We will allow λ to be time dependent in this chapter. This will turn out to be necessary for certain DAEs. We will use a different notation for stabilized differentiation. Let $\mathcal{D} = \frac{d}{dt} + \lambda$. We now apply the operator $\mathcal{D}$ to (5.1) k times to get the system

$$F = 0 \tag{5.4a}$$

$$\mathcal{D}F = 0 \tag{5.4b}$$

$$\vdots \tag{5.4c}$$

$$\mathcal{D}^k F = 0 \tag{5.4d}$$

where $F = Ax' + Bx - f$. The derivative array (5.4) can be written as

$$Jw = -\mathcal{F}x + g \tag{5.5}$$

where

$$J = \begin{bmatrix} A & 0 & 0 & 0 & \cdots & 0 \\ \mathcal{D}A + B & A & 0 & 0 & \cdots & 0 \\ \mathcal{D}^2A + 2\mathcal{D}B & 2\mathcal{D}A + B & A & 0 & \cdots & 0 \\ \mathcal{D}^3A + 3\mathcal{D}^2B & 3\mathcal{D}^2A + 3\mathcal{D}B & 3\mathcal{D}A + B & A & \cdots & 0 \\ \vdots & & & & \ddots & \vdots \end{bmatrix},$$

$$\mathcal{F} = \begin{bmatrix} B \\ \mathcal{D}B \\ \mathcal{D}^2B \\ \vdots \\ \mathcal{D}^kB \end{bmatrix}, \; g = \begin{bmatrix} f \\ \mathcal{D}f \\ \mathcal{D}^2f \\ \vdots \\ \mathcal{D}^kf \end{bmatrix}$$

with $\mathcal{D}^mX = \mathcal{D}(\mathcal{D}^{m-1}X)$ for any matrix X.

Let S, K, V be the shifting operators defined in the last chapter and let $\tilde{J}, \tilde{\mathcal{F}}, \tilde{g}$ denote the same matrices with one extra stabilized differentiation $\mathcal{D}$. We then have the following relationships

$$K\tilde{J} = JK \tag{5.6}$$

$$K\tilde{\mathcal{F}} = \mathcal{F} \tag{5.7}$$

$$K\tilde{g} = g \tag{5.8}$$

$$S\tilde{\mathcal{F}} = \mathcal{D}\mathcal{F} \tag{5.9}$$

$$S\tilde{g} = \mathcal{D}g \tag{5.10}$$

The following lemma is very important as it will be frequently used in calculations.

Lemma 4

$$S\tilde{J} = \mathcal{D}JK - \mathcal{F}V^T + JS \tag{5.11}$$

Proof. We first have

$$S\tilde{J} = \begin{bmatrix} \mathcal{D}A + B & A & 0 & 0 & \cdots & 0 \\ \mathcal{D}^2A + 2\mathcal{D}B & 2\mathcal{D}A + B & A & 0 & \cdots & 0 \\ \mathcal{D}^3A + 3\mathcal{D}^2B & 3\mathcal{D}^2A + 3\mathcal{D}B & 3\mathcal{D}A + B & A & \cdots & 0 \\ \vdots & & & & \ddots & \vdots \end{bmatrix} \tag{5.12}$$

and

$$\mathcal{D}JK = \begin{bmatrix} \mathcal{D}A & 0 & 0 & 0 & \cdots & 0 \\ \mathcal{D}^2A + \mathcal{D}B & \mathcal{D}A & 0 & 0 & \cdots & 0 \\ \mathcal{D}^3A + 2\mathcal{D}^2B & 2\mathcal{D}^2A + \mathcal{D}B & \mathcal{D}A & 0 & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \end{bmatrix} \tag{5.13}$$

The difference of which gives

$$S\tilde{J} - \mathcal{D}JK = \begin{bmatrix} -B & A & 0 & 0 & \cdots & 0 \\ -\mathcal{D}B & \mathcal{D}A + B & A & 0 & \cdots & 0 \\ -\mathcal{D}^2B & \mathcal{D}^2A + 2\mathcal{D}B & 2\mathcal{D}A + B & A & \cdots & 0 \\ \vdots & & & & \ddots & \vdots \end{bmatrix} \tag{5.14}$$

On the other hand we have

$$\mathcal{F}V^T = \begin{bmatrix} \mathcal{D}B & 0 & 0 & 0 & \cdots & 0 \\ \mathcal{D}^2B & 0 & 0 & 0 & \cdots & 0 \\ \mathcal{D}^3B & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \end{bmatrix} \tag{5.15}$$

and

$$JS = \begin{bmatrix} 0 & A & 0 & 0 & \cdots & 0 \\ 0 & \mathcal{D}A + B & A & 0 & \cdots & 0 \\ 0 & \mathcal{D}^2A + 2\mathcal{D}B & 2\mathcal{D}A + B & A & \cdots & 0 \\ \vdots & & & & \ddots & \vdots \end{bmatrix} \tag{5.16}$$

Hence the result follows. Note that $\mathcal{D}JK = J'K + \lambda JK$. □

This time we will calculate the completion in terms of Z_1, Z_2 which were defined in the previous chapter. However, we need to prove the existence of such $Z_{1,0}, Z_2$ for time variable systems as well. We will do this by using a canonical form similar to the one used in the previous chapter. The results we need are summed up in the following lemma [16], [42].

Lemma 5 *Suppose that the linear time variable DAE (5.1) is solvable, where $A(t)$ is identically singular. Then*

1. *There exists pointwise nonsingular matrices $P(t)$ and $Q(t)$ such that*

$$PAQ = \begin{bmatrix} I_d & W \\ 0 & N \end{bmatrix}, PBQ - PAQ' = \begin{bmatrix} 0 & 0 \\ 0 & I_a \end{bmatrix}$$

   In other words, left multiplication by $P(t)$ and the coordinate change given by $x = Q(t)y$ transform (5.1) to

$$\begin{bmatrix} I_d & W \\ 0 & N \end{bmatrix} y' + \begin{bmatrix} 0 & 0 \\ 0 & I_a \end{bmatrix} y = P(t)f(t) \tag{5.17}$$

   The system $Ny_2' + y_2 = f_2$ is uniquely solvable and has only one solution for sufficiently smooth f.

2. *Let $\widehat{J}_k$ be the jacobian of (5.17) obtained from k differentiations. Suppose that $\widehat{J}_k$ has constant rank. Then corank$(\widehat{J}_i) =$ dim$(G) = a$ for $i \geq k$.*

3. *Let J_k be the jacobian of (5.1) with k differentiations. Then*

$$J_k = \widehat{P}\widehat{J}_k\widehat{Q}$$

   where $\widehat{P}, \widehat{Q}$ are pointwise nonsingular matrices. Consequently, $corank(J_i) = a$ for $i \geq k$ provided that J_k has constant rank.

Note that N in (5.17) does not have to have constant rank, it is only the derivative array that needs to have constant rank. An example where N cannot be made strictly upper triangular is found in [9].

Now suppose that the derivative array (5.5) satisfies our basic assumptions (A1)–(A4) from Chapter 1 with $n = k$. Then, there exists a smooth matrix function Z_2 of full column rank satisfying $Z_2^T J = 0$. Equivalently, the columns of the matrix Z_2 form a basis for $R(J)^\perp$. Let rank$(Z_2) = a$. Since both Z_2^T and $\begin{bmatrix} J & \mathcal{F} \end{bmatrix}$ have full row rank, $Z_2^T \mathcal{F}$ also has full row rank, which will also be a. Let T_2 be a matrix function whose columns form a basis for $N(Z_2^T \mathcal{F})$. Then rank$(AT_2) = d$. let $Z_{1,0}$ be a matrix whose d columns form a basis for $R(AT_2)$. We have $a + d = n$. Lets define $Z_1^T = \begin{bmatrix} Z_{1,0}^T & 0 & \cdots & 0 \end{bmatrix}$.

Lemma 6 *The matrix function* $\begin{bmatrix} Z_{1,0}^T A \\ Z_2^T \mathcal{F} \end{bmatrix}$ *is pointwise invertible.*

Proof. Let $t = t_0$. Suppose that

$$\begin{bmatrix} (Z_{1,0}^T A)(t_0) \\ (Z_2^T \mathcal{F})(t_0) \end{bmatrix} v = 0$$

for some vector v. Then, we have $(Z_{1,0}^T A)(t_0)v = 0$ and $(Z_2^T \mathcal{F})(t_0)v = 0$. $(Z_2^T \mathcal{F})(t_0)v = 0$ implies $v \in N(Z_2^T \mathcal{F}(t_0)) = R(T_2(t_0))$ by definition. Therefore, $v = T_2 x$ for some vector x. Then, by the first part, we get $(Z_{1,0}^T AT_2)(t_0)x = 0 = Z_{1,0}^T (AT_2)(t_0)x$, which implies $(AT_2)(t_0)x = 0$ by definition of $Z_{1,0}$. But $(AT_2)(t_0)$ has full column rank, thus we get $x = 0$, so $v = T_2 x = 0$. □

5.3 Calculating the Completion

We will now calculate the LSC by finding a matrix G_0 that satisfies the two conditions

$$G_0 \tilde{J} = \begin{bmatrix} I & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{5.18a}$$

$$G_0 \tilde{Z}_2 = 0 \tag{5.18b}$$

where $\tilde{Z}_2$ is a matrix of maximal rank satisfying $\tilde{Z}_2^T \tilde{J} = 0$. Note that the LSC is then given by the formula

$$x' = G_0(-\tilde{\mathcal{F}} x + \tilde{g}). \tag{5.19}$$

Consider the following matrix

$$G_0 = D \begin{bmatrix} Z_1^T K + C_1 Z_2^T K \\ (Z_2^T)' K + Z_2^T S + C_2 Z_2^T K \end{bmatrix} \tag{5.20}$$

We claim that suitable matrices C_1, C_2 and D exist such that G_0 satisfies the above conditions. Note that since corank$(\tilde{J})$ = corank(J), $\widehat{Z}_2^T K$ is a matrix of maximal rank satisfying $(Z_2^T K)\tilde{J} = 0$. Therefore, we can choose $\tilde{Z}_2^T = Z_2^T K$.

Using Lemma 4 and the fact that $Z_2^T J = 0$, we have

$$\begin{aligned}
(Z_1^T K + C_1 Z_2^T K)\tilde{J} &= Z_1^T K \tilde{J} + C_1 Z_2^T K \tilde{J} \\
&= Z_1^T J K + C_1 Z_2^T J K \\
&= Z_1^T J K \\
&= \begin{bmatrix} Z_{1,0}^T A & 0 & 0 & \cdots & 0 \end{bmatrix}
\end{aligned} \tag{5.21}$$

and

$$\begin{aligned}
((Z_2^T)' K + Z_2^T S + C_2 Z_2^T K)\tilde{J} &= (Z_2^T)' K \tilde{J} + Z_2^T (J' K + \lambda J K - \mathcal{F} V^T + J S) + C_2 Z_2^T K \tilde{J} \\
&= (Z_2^T)' K \tilde{J} + Z_2^T J' K - Z_2^T \mathcal{F} V^T \\
&= (Z_2^T)' J K + Z_2^T J' K - Z_2^T \mathcal{F} V^T \\
&= (Z_2^T J)' K - Z_2^T \mathcal{F} V^T \\
&= -Z_2^T \mathcal{F} V^T = \begin{bmatrix} -Z_2^T \mathcal{F} & 0 & 0 & \cdots & 0 \end{bmatrix}
\end{aligned} \tag{5.22}$$

Thus we have

$$G_0 \tilde{J} = D \begin{bmatrix} Z_1^T K + C_1 Z_2^T K \\ (Z_2^T)' K + Z_2^T S + C_2 Z_2^T K \end{bmatrix} \tilde{J} = D \begin{bmatrix} Z_{1,0}^T A & 0 & \cdots & 0 \\ -Z_2^T \mathcal{F} & 0 & \cdots & 0 \end{bmatrix} \tag{5.23}$$

Since $\begin{bmatrix} Z_{1,0}^T A \\ -Z_2^T \mathcal{F} \end{bmatrix}$ is pointwise invertible, the choice $D = \begin{bmatrix} Z_{1,0}^T A \\ -Z_2^T \mathcal{F} \end{bmatrix}^{-1}$ will give us the first condition

$$G \tilde{J} = \begin{bmatrix} I & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{5.24}$$

We will use the condition (5.18b) now to find C_1, C_2. So suppose that $G_0\tilde{Z}_2 = 0$. Then,

$$(Z_1^T K + C_1 Z_2^T K)\tilde{Z}_2 = 0, \tag{5.25a}$$

$$((Z_2^T)' K + Z_2^T S + C_2 Z_2^T K)\tilde{Z}_2 = 0 \tag{5.25b}$$

From this we get

$$\begin{aligned} C_1 &= -(Z_1^T K \tilde{Z}_2)(Z_2^T K (Z_2^T K)^T)^{-1} \\ &= -(Z_1^T K (Z_2^T K)^T)(Z_2^T K (Z_2^T K)^T)^{-1} \\ &= -(Z_1^T Z_2)(Z_2^T Z_2)^{-1} \end{aligned} \tag{5.26}$$

and

$$\begin{aligned} C_2 &= -((Z_2^T)' K + Z_2^T S)(Z_2^T K)^T (Z_2^T K (Z_2^T K)^T)^{-1} \\ &= -(Z_2^T)' Z_2 (Z_2^T Z_2)^{-1} + Z_2^T S K^T Z_2 (Z_2^T Z_2)^{-1} \end{aligned} \tag{5.27}$$

Therefore, G_0 is the matrix we are looking for with C_1, C_2, D as defined above. Thus, we can now calculate the LSC by the formula

$$x' = G_0(-\tilde{\mathcal{F}} x + \tilde{g}) \tag{5.28}$$

Then we have

$$\begin{aligned} (Z_1^T K + C_1 Z_2^T K)\tilde{\mathcal{F}} &= Z_1^T \mathcal{F} + C_1 Z_2^T \mathcal{F} \\ &= Z_{1,0}^T B + C_1 (Z_2^T \mathcal{F}) \end{aligned} \tag{5.29}$$

$$\begin{aligned} (Z_1^T K + C_1 Z_2^T K)\tilde{g} &= Z_1^T g + C_1 Z_2^T g \\ &= Z_{1,0}^T f + C_1 (Z_2^T g) \end{aligned} \tag{5.30}$$

and

$$\begin{aligned}
((Z_2^T)'K + Z_2^T S + C_2 Z_2^T K)\tilde{\mathcal{F}} &= (Z_2^T)'K\tilde{\mathcal{F}} + Z_2^T S\tilde{\mathcal{F}} + C_2 Z_2^T K\tilde{\mathcal{F}} \\
&= (Z_2^T)'\mathcal{F} + Z_2^T(\mathcal{D}\mathcal{F}) + C_2 Z_2^T\mathcal{F} \\
&= (Z_2^T)'\mathcal{F} + Z_2^T(\mathcal{F}' + \lambda\mathcal{F}) + C_2 Z_2^T\mathcal{F} \\
&= (C_2 + \lambda I)(Z_2^T\mathcal{F}) + (Z_2^T\mathcal{F})' \qquad (5.31)
\end{aligned}$$

Similarly

$$((Z_2^T)'K + Z_2^T S + C_2 Z_2^T K)\tilde{g} = (C_2 + \lambda I)(Z_2^T g) + (Z_2^T g)' \qquad (5.32)$$

Therefore the LSC is given by

$$x' = G(-\tilde{\mathcal{F}} + \tilde{g})x = D\begin{bmatrix} Z_1^T K + C_1 Z_2^T K \\ (Z_2^T)'K + Z_2^T S + C_2 Z_2^T K \end{bmatrix}(-\tilde{\mathcal{F}}x + \tilde{g})$$

which becomes

$$x' = \begin{bmatrix} Z_{1,0}^T A \\ -Z_2^T\mathcal{F} \end{bmatrix}^{-1}\left(-\begin{bmatrix} C_1(Z_2^T\mathcal{F}) + Z_{1,0}^T B \\ (C_2 + \lambda I)(Z_2^T\mathcal{F}) + (Z_2^T\mathcal{F})' \end{bmatrix}x + \begin{bmatrix} C_1(Z_2^T g) + Z_{1,0}^T f \\ (C_2 + \lambda I)(Z_2^T g) + (Z_2^T f)' \end{bmatrix}\right) \qquad (5.33)$$

Note that the formula contains the term $(Z_2^T\mathcal{F})'$ which is equal to $Z_2'^T\mathcal{F} + Z_2^T\mathcal{F}'$ since Z_2 is now time variable. Therefore the formula requires the calculation of Z_2'. However, as commented on earlier, since Z_2 is calculated numerically, computing Z_2' directly by differentiating Z_2 can cause too much numerical error. In Chapter 4 we outlined a process on how to calculate the alternative completion using the least squares method. In Chapter 6 we will modify the same process for LTV systems and show how that can be used to calculate the completion without having to differentiate Z_2.

Note that the completion (5.33) can also be written as

$$Z_{1,0}^T(Ax' + Bx - f) = -C_1(Z_2^T\mathcal{F}x - Z_2^T g) \qquad (5.34a)$$

$$(Z_2^T\mathcal{F}x - Z_2^T g)' = -(C_2 + \lambda I)(Z_2^T\mathcal{F}x - Z_2^T g) \qquad (5.34b)$$

This will be the main form we will use for our analysis.

5.4 The Uniqueness

Note that Z_1 and Z_2 are not uniquely determined. However, since the LSC is unique, this formula should not depend on the choice of Z_1, Z_2. The following lemma is to confirm this fact.

Lemma 7 *The LSC given by (5.34) is independent of which $Z_{1,0}, Z_2$ are taken.*

Proof. Suppose that $\widehat{Z}_{1,0}, \widehat{Z}_2$ are two matrices with the same property as $Z_{1,0}, Z_2$. Since their columns constitute a basis for the same subspaces, there exists nonsingular matrices $P(t)$ and $Q(t)$ such that

$$\begin{aligned} \widehat{Z}_{1,0}^T &= PZ_{1,0}^T, & (5.35)\\ \widehat{Z}_2^T &= QZ_2^T & (5.36) \end{aligned}$$

P, Q are smooth if the Z matrices are. Then we get

$$\begin{aligned} \widehat{Z}_1^T &= \begin{bmatrix} PZ_{1,0}^T & 0 & \cdots & 0 \end{bmatrix} \\ &= PZ_1^T \end{aligned} \tag{5.37}$$

$$\begin{aligned} \widehat{C}_1 &= -\widehat{Z}_1^T \widehat{Z}_2 (\widehat{Z}_2^T \widehat{Z}_2)^{-1} \\ &= (-PZ_1^T)(Z_2Q^T)(QZ_2^T Z_2 Q^T)^{-1} \\ &= -PZ_1^T Z_2 (Z_2^T Z_2)^{-1} Q^{-1} \\ &= PC_1Q^{-1} \end{aligned} \tag{5.38}$$

$$\begin{aligned} \widehat{C}_2 &= -(QZ_2^T)'(Z_2Q^T)(QZ_2^TZ_2Q^T)^{-1} + (QZ_2^T)SK^T(Z_2Q^T)(QZ_2^TZ_2Q^T)^{-1} \\ &= -(Q'Z_2^T + Q(Z_2^T)')Z_2(Z_2^TZ_2)^{-1}Q^{-1} + QZ_2^TSK^TZ_2(Z_2^TZ_2)^{-1}Q^{-1} \\ &= -Q'Z_2^TZ_2(Z_2^TZ_2)^{-1}Q^{-1} - Q(Z_2^T)'Z_2(Z_2^TZ_2)^{-1}Q^{-1} + QZ_2^TSK^TZ_2(Z_2^TZ_2)^{-1}Q^{-1} \\ &= -Q'Q^{-1} - Q[(Z_2^T)'Z_2(Z_2^TZ_2)^{-1} + Z_2^TSK^TZ_2(Z_2^TZ_2)^{-1}]Q^{-1} \\ &= -Q'Q^{-1} + QC_2Q^{-1} \end{aligned} \tag{5.39}$$

Now substituting (5.38) and (5.39) in (5.34), we get

$$\widehat{Z}_{1,0}^T(Ax' + Bx - f) = \widehat{C}_1(\widehat{Z}_2^T\mathcal{F}x - \widehat{Z}_2^Tg) \tag{5.40a}$$
$$(\widehat{Z}_2^T\mathcal{F}x - \widehat{Z}_2^Tg)' = -(\widehat{C}_2 + \lambda I)(\widehat{Z}_2^T\mathcal{F}x - \widehat{Z}_2^Tg) \tag{5.40b}$$

which implies

$$PZ_{1,0}^T(Ax' + Bx - f) = PC_1Q^{-1}(QZ_2^T\mathcal{F}x - QZ_2^Tg) \tag{5.41a}$$
$$(QZ_2^T\mathcal{F}x - QZ_2^Tg)' = -(-Q'Q^{-1} + QC_2Q^{-1} + \lambda I)(QZ_2^T\mathcal{F}x - QZ_2^Tg) \tag{5.41b}$$

or

$$PZ_{1,0}^T(Ax' + Bx - f) = PC_1(Z_2^T\mathcal{F}x - Z_2^Tg) \tag{5.42a}$$
$$\begin{aligned} Q'(Z_2^T\mathcal{F}x - Z_2^Tg) + Q(Z_2^T\mathcal{F}x - Z_2^Tg)' &= (Q'Q^{-1}Q(Z_2^T\mathcal{F}x - Z_2^Tg) \\ &- (QC_2Q^{-1} + \lambda I)Q(Z_2^T\mathcal{F}x - Z_2^Tg) \end{aligned} \tag{5.42b}$$

This gives

$$PZ_{1,0}^T(Ax' + Bx - f) = PC_1(Z_2^T\mathcal{F}x - Z_2^Tg) \tag{5.43a}$$
$$Q(Z_2^T\mathcal{F}x - Z_2^Tg)' = -Q(C_2 + \lambda I)(Z_2^T\mathcal{F}x - Z_2^Tg) \tag{5.43b}$$

and finally

$$Z_{1,0}^T(Ax' + Bx - f) = C_1(Z_2^T\mathcal{F}x - Z_2^Tg) \tag{5.44a}$$
$$(Z_2^T\mathcal{F}x - Z_2^Tg)' = -(C_2 + \lambda I)(Z_2^T\mathcal{F}x - Z_2^Tg) \tag{5.44b}$$

since P and Q are pointwise nonsingular. □

Now that the completion is independent of which Z_1, Z_2 are used, we can select them in a way that will facilitate our analysis.

Lemma 8 *There exists appropriate Z_1 and Z_2 with $\|Z_1\| = \|Z_2\| = 1$ such that*

$$\|C_2\| \le 1, \ \|C_2\| \le 1 \tag{5.45}$$

Proof. We can choose a Z_2 to have orthonormal columns so that

$$Z_2^TZ_2 = I_n.$$

(If Z_2^T does not have this property, then by Gram-Schmidt, there exists a pointwise nonsingular R such that RZ_2^T does.) Now, let $Q(t)$ be the solution of the ODE

$$X' = -X(Z_2^T)'Z_2, X(t_0) = I$$

Then $Q(t)$ is smooth and pointwise nonsingular. Moreover we have

$$\begin{aligned}
(QQ^T)' &= Q'Q^T + QQ'^T \\
&= -(Q(Z_2^T)'Z_2)Q^T - Q(Q(Z_2^T)'Z_2)^T \\
&= -Q(Z_2^T)'Z_2Q^T - QZ_2^T(Z_2^T)'^TQ^T \\
&= -Q((Z_2^T)'Z_2 + Z_2^T(Z_2')Q^T \\
&= -Q(Z_2^TZ_2)'Q^T = 0 \qquad (5.46)
\end{aligned}$$

since $Z_2^TZ_2 = I$. Therefore, we have $QQ^T = I$ and $Q(t)$ is point-wise unitary by the initial condition [42]. Let $\widehat{Z}_2^T = QZ_2^T$. Then, using $Z_2^TZ_2 = I = QQ^T$, we get

$$\begin{aligned}
H_1 &= (\widehat{Z}_2^T)'\widehat{Z}_2(\widehat{Z}_2^T\widehat{Z}_2)^{-1} \\
&= (QZ_2^T)'(QZ_2^T)^T(QZ_2^TZ_2Q^T)^{-1} \\
&= (Q'Z_2^T + Q(Z_2^T)')Z_2Q^T(I) \\
&= (-Q(Z_2^T)'Z_2Z_2^T + Q(Z_2^T)')Z_2Q^T \\
&= -Q(Z_2^T)'Z_2Z_2^TZ_2Q^T + Q(Z_2^T)'Z_2Q^T \\
&= -Q(Z_2^T)'Z_2Q^T + Q(Z_2^T)'Z_2Q^T = 0 \qquad (5.47)
\end{aligned}$$

and

$$\begin{aligned}
H_2 &= (QZ_2^T)SK^T(QZ_2^T)^T((QZ_2^T)(QZ_2^T)^T)^{-1} \\
&= (QZ_2^T)SK^T(QZ_2^T)^T \\
&= QZ_2^TSK^TZ_2Q^T \qquad (5.48)
\end{aligned}$$

Let $\|,\|$ be the spectral norm. Then,

$$\begin{aligned} \|\widehat{C}_2\| &= \|\widehat{H}_1 + \widehat{H}_2\| \\ &\leq \|\widehat{H}_1\| + \|\widehat{H}_2\| \\ &= 0 + \|QZ_2^T SK^T Z^T Q^T\| \\ &\leq \|Q\|\|Z_2^T\|\|S\|\|K^T\|\|Z_2\|\|Q^T\| \\ &= 1 \end{aligned} \tag{5.49}$$

since $\|Q\| = \|Z_2^T\| = \|S\| = \|K\| = 1$. □

Similarly, let $Z_{1,0}^T$ have an orthonormal set of rows and Z_2 be as calculated above. Then, $Z_1^T Z_1 = I$ and

$$\|\widehat{C}_1\| = \|\widehat{Z}_1^T \widehat{Z}_2^T (\widehat{Z}_2^T \widehat{Z}_2)^{-1}\| \leq 1 \tag{5.50}$$

5.5 The Stability Properties

We will now begin analyzing the additional dynamics of the completion (5.34), which involves the parameter $\lambda(t)$. We want to examine how λ effects the stability of the system and determine the sufficient conditions that λ needs to satisfy in order to achieve a desired stability. Instead of determining the additional dynamics explicitly, as done in the constant coefficient case, we will directly look at the behavior of the solutions of the completion relative to the solution manifold. The convergence of an arbitrary solution of the completion to the solution manifold is what we are looking for. Throughout this thesis it is assumed that $\lambda(t) > 0$.

Let σ denote the smallest singular value of a matrix and let

$$\delta(t) = \sigma(D^{-1}) = \sigma\left(\begin{bmatrix} Z_{1,0}^T A \\ Z_2^T \mathcal{F} \end{bmatrix}\right) \tag{5.51}$$

We have $\delta(t) > 0$ for all $t > 0$ but $\delta(t)$ is not necessarily bounded away from zero.

We are now in a position to give our first stabilization result.

Theorem 9 *Let $M(t) = \{a \in \mathbb{R}^n | Z_2^T \mathcal{F}(t)a - Z_2^T g(t) = 0\}$. That is, $M = \{M(t)|t \in \mathbb{R}\}$ is the solution manifold of the DAE. Let $y = y(t)$ be an arbitrary solution of the LSC*

(5.34). Suppose that we have a smooth k such that

$$e^{-k(t)} \le \delta(t) \tag{5.52}$$

Let $\overline{\lambda}(t) = \int_0^t \lambda(\tau)d\tau$. *Then*

$$d(y(t), M(t)) \le c_0 e^{-\overline{\lambda}(t)+t+k(t)} \tag{5.53}$$

Here d *is the standard metric, and* $d(y(t), M(t)) = \inf\{d(y(t), a)|a \in M(t)\}$. *In particular, if* $\int_0^\infty -\lambda(\tau) + 1 + k'(\tau)d\tau = -\infty$, *then*

$$lim_{t\mapsto\infty} d(y(t), M(t)) = 0 \tag{5.54}$$

Proof. Let $y = y(t)$ be an arbitrary solution of the LSC (5.34). Let $x = \begin{bmatrix} Z_{1,0}^T A \\ Z_2^T \mathcal{F} \end{bmatrix}^{-1} \begin{bmatrix} Z_{1,0}^T Ay \\ Z_2^T g \end{bmatrix}$ so that we have

$$\begin{aligned} Z_{1,0}^T Ax &= Z_{1,0}^T Ay && \text{(5.55a)} \\ Z_2^T \mathcal{F}x - Z_2^T g &= 0 && \text{(5.55b)} \end{aligned}$$

Then (5.55b) implies $x(t) \in M(t)$. Note that we are not saying at this point that $x(t)$ is a solution of the DAE, merely that at each t it is a consistent initial condition. On the other hand, by (5.34b) $Z_2^T \mathcal{F}y - Z_2^T g$ is a solution of the ODE

$$z' = -(C_2 + \lambda I)z$$

Then, there is a constant a_0 so that $\|z\| \le a_0 e^{\int_0^t \|C_2(\tau)\| - \lambda(\tau)d\tau}$. Therefore, by (5.49), we get

$$\|(Z_2^T \mathcal{F}y - Z_2^T g)(t)\| \le a_0 e^{-\overline{\lambda}(t)+t} \tag{5.56}$$

where $\overline{\lambda}(t) = \int_0^t \lambda(\tau)d\tau$. Combining this with (5.55b), we get for $t \ge 0$ that

$$\|(Z_2^T \mathcal{F}y - Z_2^T g) - (Z_2^T \mathcal{F}x - Z_2^T g)\| \le a_0 e^{-\overline{\lambda}(t)+t} \tag{5.57}$$

and hence

$$\|Z_2^T \mathcal{F}(y - x)\| \le a_0 e^{-\overline{\lambda}(t)+t} \tag{5.58}$$

On the other hand, by (5.55a) we have

$$Z_{1,0}^T A(y-x)(t) = 0 \tag{5.59}$$

Now combining (5.58) and (5.59), we get

$$\| \begin{bmatrix} Z_{1,0}^T A \\ -Z_2^T \mathcal{F} \end{bmatrix} (y-x)(t)\| = \|D^{-1}(y-x)(t)\| \leq a_0 e^{-\overline{\lambda}(t)+t} \tag{5.60}$$

But by definition,

$$\|D^{-1}(y-x)\| \geq \delta(t)\|y-x\| \tag{5.61}$$

Then (5.61), (5.60), (5.52) imply

$$\|y(t)-x(t)\| \leq c_0 e^{-\overline{\lambda}(t)+t+k(t)} \tag{5.62}$$

Theorem (9) now follows. □

Example 4 *Consider the pure algebraic DAE*

$$e^{-2t}x = f(t) \tag{5.63}$$

where f is a differentiable function. Note that the solution manifold for this DAE is $x(t) = e^{2t}f(t)$. Lets now calculate the stabilized LSC. One differentiation in the stabilized sense gives

$$\begin{aligned} 0 &= e^{-2t}x - f && (5.64)\\ 0 &= e^{-2t}x' - 2e^{-2t}x + \lambda e^{-2t}x - f' - \lambda f && (5.65)\\ &= e^{-2t}x' + (\lambda-2)e^{-2t}x - f' - \lambda f && (5.66) \end{aligned}$$

Therefore the LSE is

$$\begin{bmatrix} 0 & 0 \\ e^{-2t} & 0 \end{bmatrix}^T \begin{bmatrix} 0 & 0 \\ e^{-2t} & 0 \end{bmatrix} \begin{bmatrix} x' \\ x'' \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ e^{-2t} & 0 \end{bmatrix}^T \left(- \begin{bmatrix} e^{-2t} \\ (\lambda-2)e^{-2t} \end{bmatrix} x + \begin{bmatrix} f \\ f' + \lambda f \end{bmatrix} \right)$$

which simplifies to

$$\begin{bmatrix} e^{-4t} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x' \\ x'' \end{bmatrix} = - \begin{bmatrix} (\lambda-2)e^{-4t} \\ 0 \end{bmatrix} x + \begin{bmatrix} e^{-2t}(f' + \lambda f) \\ 0 \end{bmatrix}$$

The first block row of (4) gives us the LSC as

$$x' = -(\lambda - 2)x + e^{2t}(f' + \lambda f) \tag{5.67}$$

The difference ϵ *between* x *of (5.67) and the unique solution* $x = e^{2t}f$ *satisfies* $\epsilon' = -(\lambda-2)\epsilon$ *and* ϵ *goes to zero if* λ *is large enough. In particular, we need* $\int_0^\infty (2-\lambda)d\tau = -\infty$.

Example 5 *Now consider*

$$\begin{aligned} x_1' &= -x_1 + e^{\alpha t}x_2 && (5.68a) \\ 0 &= x_2 && (5.68b) \end{aligned}$$

whose stabilized LSC is

$$\begin{aligned} x_1' &= -x_1 + e^{\alpha t}x_2 && (5.69a) \\ x_2' &= -\lambda x_2 && (5.69b) \end{aligned}$$

Suppose that λ *is constant and* $\alpha \geq 0$. *The solutions of (5.68) are* $x^T = [e^{-t}c, 0]$ *so that* $M(t)$ *is just the constant subspace spanned by* $[1, 0]^T$. *On the other hand, the solutions of (5.69) are*

$$x_2 = e^{-\lambda t}c_2, \; x_1 = \begin{cases} c_1e^{-t} + c_2\frac{1}{1+\alpha-\lambda}e^{(\alpha-\lambda)t} & \text{if } 1+\alpha-\lambda \neq 0 \\ c_1e^{-t} + c_2te^{-t} & \text{if } 1+\alpha-\lambda = 0 \end{cases} \tag{5.70}$$

We see that $\lim_{t\to\infty} d(y(t), M(t)) = 0$ *for every solution of the completion if and only if* $\lambda > 0$. *However, for every solution of the completion there is a solution* x *of the DAE so that* $\lim_{t\to\infty}(y(t) - x(t)) = 0$ *if and only if* $\lambda > \alpha$.

Theorem 9 states that the solutions of the completion will converge to the solution manifold for appropriately chosen λ's. However, this does not necessarily mean that they will converge to a particular solution of the DAE as well. As the Example 5 shows, this might require an even larger λ in general. Theorem 10 describes the conditions under which such a convergence can take place. We have two separate results. First, given a solution $y(t)$ of the completion with the initial condition $y(0)$, if we take the solution $x(t)$ of the DAE that is closest to $y(t)$ at the time $t = 0$, then the difference $\|y(t) - x(t)\|$ is bounded by a function involving λ. Thus, it is possible to lower the bound by choosing λ appropriately. However, for any given λ, the bounding function is still monotone increasing, therefore, $y(t)$ may still move away from $x(t)$, and thus the convergence at infinity would not occur. Therefore, this result is most useful when a finite interval is concerned.

On the other hand, there is also a solution of the DAE, $\widehat{x}(t)$, while it is not the closest one to $y(t)$ at $t = 0$, for which we do get $\|y(t) - \widehat{x}(t)\| \to 0$ provided that λ satisfies certain conditions.

Now, note that the system

$$Z_{1,0}^T(Ax' + Bx - f) = 0 \tag{5.71a}$$

$$Z_2^T \mathcal{F}x - Z_2^T g = 0 \tag{5.71b}$$

has the same solutions as the original DAE. On the other hand, we have calculated the LSC as

$$Z_{1,0}^T(Ax' + Bx - f) = -C_1(Z_2^T \mathcal{F}x - Z_2^T g) \tag{5.72a}$$

$$(Z_2^T \mathcal{F}x - Z_2^T g)' = -(C_2 + \lambda I)(Z_2^T \mathcal{F}x - Z_2^T g) \tag{5.72b}$$

Therefore, the difference $z = z(t) = y(t) - x(t)$ between a solution $y = y(t)$ of (5.72) and a solution $x = x(t)$ of the DAE, which will satisfy (5.71), satisfies the equation

$$Z_{1,0}^T(Az' + Bz) = -C_1(Z_2^T \mathcal{F}z) \tag{5.73a}$$

$$(Z_2^T \mathcal{F}z)' = -(C_2 + \lambda I)(Z_2^T \mathcal{F}z) \tag{5.73b}$$

Lets define the change of variable

$$\tilde{z} = \begin{bmatrix} \tilde{z}_1 \\ \tilde{z}_2 \end{bmatrix} = \begin{bmatrix} Z_{1,0}^T A \\ -Z_2^T \mathcal{F} \end{bmatrix} z = D^{-1} z$$

so that we have

$$\tilde{z}_1 = Z_{10}^T A z, \tag{5.74a}$$

$$\tilde{z}_2 = -Z_2^T \mathcal{F} z \tag{5.74b}$$

Then, (5.73) becomes

$$\tilde{z}_1' + U_1 \tilde{z}_1 = -U_2 \tilde{z}_2 + C_1 \tilde{z}_2 \tag{5.75a}$$

$$\tilde{z}_2' = -(C_2 + \lambda I)\tilde{z}_2 \tag{5.75b}$$

where

$$U = ((Z_{1,0}^T A)' + Z_{1,0}^T B)D \tag{5.76a}$$

$$U_1 = U[I, 0]^T \tag{5.76b}$$

$$U_2 = U[0, I]^T \tag{5.76c}$$

Note that we have

$$\|U_1(t)\| \leq \|U(t)\|, \text{ and } \|U_2(t)\| \leq \|U(t)\| \tag{5.77}$$

Let $U_t = sup\{\|U(s)\| \,|\, s \in [0, t]\}$. Then, U_t is monotone increasing. For the sake of simplicity, we will also assume that $k(t)$ is nondecreasing as well, which is not against the nature of $k(t)$.

We can now prove our second result.

Theorem 10 *Let $y = y(t)$ be a solution of the LSC (5.34) with the initial condition $y(0) = y_0$. Let $x = x(t)$ be the solution of the DAE with the initial condition*

$$x(0) = \begin{bmatrix} Z_{1,0}^T A \\ -Z_2^T \mathcal{F} \end{bmatrix}^{-1} \begin{bmatrix} Z_{1,0}^T A y \\ -Z_2^T g \end{bmatrix} (0) \tag{5.78}$$

Then we have

$$\|y(t) - x(t)\| \leq b_0 \lambda_1(t) \tag{5.79}$$

where $\lambda_1(t) = (U_t t + t)e^{2U_t t + t + k(t)} \int_0^t e^{-\overline{\lambda}(\tau)} d\tau$*. In particular, if* λ *is such that* $\lambda_1'(t) \leq e^{-t}$*, then there exists another solution* $\widehat{x} = \widehat{x}(t)$ *of the DAE such that*

$$\lim_{t\to\infty} \|y(t) - \widehat{x}(t)\| = 0 \tag{5.80}$$

Proof. By the same argument as in the previous theorem, (5.75b) implies

$$\|\tilde{z}_2(t)\| \leq a_0 e^{-\overline{\lambda}(t)+t} \tag{5.81}$$

On the other hand, (5.75a) implies

$$z_1(t) = \Phi(t)\left(\Phi^{-1}(0)z_1(0) + \int_0^t \Phi^{-1}(\tau)R(\tau)d\tau\right) \tag{5.82}$$

where $\Phi(t)$ is a fundamental matrix solution of $z_1' = -U_1 z_1$ and $R(t) = (-U_2 + C_1)\tilde{z}_2(t)$. Let $R_t = \sup\{\|R(s)\| \,|\, s \in [0,t]\}$. Note that R_t is also positive increasing. Then, (5.50) and (5.81) imply

$$R_t \leq a_0(U_t + 1)e^{-\overline{\lambda}(t)+t} \tag{5.83}$$

On the other hand, since $\Phi(t)$ is a fundamental matrix solution of $z' = -U_1 z$, we have $\Phi(t) = -\int_0^t U_1(\tau)\Phi(\tau)d\tau$. By (5.77) this implies

$$\|\Phi(t)\| \leq U_t \int_0^t \|\Phi(\tau)\| d\tau \tag{5.84}$$

Thus, by Gronwall,

$$\|\Phi(t)\| \leq e^{U_t t} \tag{5.85}$$

Similarly, $\Phi^{-1}(t)$ is a fundamental matrix solution for $z' = zU_1$. Therefore, by the same argument, we have

$$\|\Phi^{-1}(t)\| \leq e^{U_t t} \tag{5.86}$$

Note also that by the assumption (5.78) we have $z_1(0) = \tilde{y}_1(0) - \tilde{x}_1(0) = Z_{1,0}Ay(0) - Z_{1,0}Ax(0) = 0$. Therefore, substituting (5.83), (5.85), (5.86) in (5.82), we get

$$\begin{aligned} \|\tilde{z}_1(t)\| &\leq a_0 e^{U_t t} \int_0^t e^{U_\tau \tau}(U_\tau + 1)e^{-\overline{\lambda}(\tau)+\tau} d\tau \\ &\leq a_0(U_t t + t)e^{2U_t t + t} \int_0^t e^{-\overline{\lambda}(\tau)} d\tau \end{aligned} \tag{5.87}$$

since U_t and k are increasing. Combining that with (5.81) we get

$$\|\tilde{z}(t)\| = \|(D^{-1}z)(t)\| \leq a_0(U_t t + t)e^{2U_t t + t} \int_0^t e^{-\overline{\lambda}(\tau)} d\tau \tag{5.88}$$

Thus, by (5.52) we arrive at

$$\|z(t)\| \leq b_0(U_t t + t)e^{2U_t t + t + k(t)} \int_0^t e^{-\overline{\lambda}(\tau)} d\tau = b_0 \lambda_1(t) \tag{5.89}$$

which gives us the first part of theorem.

To prove the second part, suppose that $\lambda_1'(t) \leq e^{-t}$. Then by (5.89) this implies

$$(U_t + 1)e^{-\overline{\lambda}(t) + 2U_t t + t + k(t)} \leq e^{-t} \tag{5.90}$$

for all $t \geq 0$. Then

$$\|\Phi^{-1}(s)R(s)\| \leq (U_t + 1)e^{-\overline{\lambda}(t) + t + U_t t} \leq e^{-U_t t - t - k(t)} \tag{5.91}$$

Thus the integral $\int_0^\infty \Phi^{-1}(\tau)R(\tau)d\tau$ is defined. Let

$$z_1(0) = -(\Phi^{-1}(0))^{-1} \int_0^\infty \Phi^{-1}(\tau)R(\tau)d\tau \tag{5.92}$$

Then substituting (5.92), (5.91) and (5.85) in (5.82) we get

$$\begin{aligned} \|\tilde{z}_1(t)\| \leq e^{U_t t}\left(\int_t^\infty e^{-U_\tau \tau - \tau - k(t)} d\tau\right) &\leq e^{U_t t}\left(e^{-U_t t - k(t)} \int_t^\infty e^{-\tau} d\tau\right) \\ &= e^{U_t t} e^{-U_t t - k(t)} e^{-t} = e^{-t - k(t)} \end{aligned} \tag{5.93}$$

Combining (5.93) with (5.81) and (5.90), we get

$$\|\tilde{z}(t)\| \leq c_0 e^{-t - k(t)} \tag{5.94}$$

Again, since $\|\tilde{z}(t)\| = \|(D^{-1}z)(t)\|$, this implies

$$\|z(t)\| \leq d_0 e^{-t - k(t) + k(t)} = d_0 e^{-t} \tag{5.95}$$

Thus

$$\lim_{t \to \infty} \|z(t)\| = \lim_{t \to \infty} \|y(t) - x(t)\| = 0$$

□

Note that $x(t)$ is a solution of the DAE that is specified by the initial condition

$$z_0 = -(\Phi^{-1}(0))^{-1} \int_0^{\infty} \Phi^{-1}(\tau) R(\tau) d\tau \tag{5.96}$$

There is no restrictions on $\tilde{z}_2(0)$ however. Since $\tilde{z}(0) = D^{-1}z(0) = D^{-1}(x_0 - y_0)$, (5.96) is equivalent to $x(0) = D\tilde{z}(0) + y(0)$. Therefore, the solution of the completion with the initial value $y(0) = (y_1(0), y_2(0))$ will converge to solutions of the DAE with initial value

$$\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = D \begin{bmatrix} \tilde{z}_1(0) \\ c \end{bmatrix} + \begin{bmatrix} y_1(0) \\ y_2(0) \end{bmatrix} \tag{5.97}$$

where c is an arbitrary constant vector. Therefore, in fact, each solution of the completion will converge to a family of solutions of the DAE.

Remark 1 *Note that since J depends on λ, the matrices $Z_{1,0}, Z_2$ and thus D^{-1}, U, C_1 and C_2 can depend on λ as well. Therefore, Theorem 9 and Theorem 10 do not necessarily imply the convergence for a large enough $\lambda = \lambda(t)$.*

In the constant coefficient case, we had proved that the additional dynamics consisted of a polynomial times an exponential, and the manifold was asymptotically stable for any positive real λ. However, as we have just showed, this is not the case for the time variable systems. λ has to satisfy certain assumptions. Also, even when the stability is obtained, we don't have a clear description of the additional dynamics as in the constant coefficient case. Moreover, selecting an appropriate λ that will provide the stability can involve technical difficulties since it depends on the matrices such as $D(t)$ and $U(t)$. These difficulties motivate us to consider the alternative stabilized completion for LTV systems, which we will begin to analyze now.

Chapter 6

Alternative Stabilized Completion: LTV Systems

6.1 Index One Formulation and Stability

In Chapter 4 we have introduced the alternative stabilized completion for constant coefficient systems. We will now analyze the extension of that method to time variable DAEs. While the alternative stabilized completion is computationally more expensive than the stabilized LSC, we will show that it has better stability properties for time variable systems and can overcome some of the difficulties present with the first technique. The extension process will basically consist of replacing the elements described in Chapter 4 with their time variable versions. However, the smoothness of the components will be an issue as well for time variable DAEs.

Let $Z_{1,0}, Z_2$ be the matrices defined in the previous chapter, but corresponding to the standard derivative array now. Thus, they don't depend on λ this time. We have already proved existence of smooth $Z_{1,0}, Z_2$ for the derivative array (5.4), and the standard derivative array is a special case of (5.4) with $\mathcal{D} = \frac{d}{dt}$. Therefore, we have that the system

$$Z_{1,0}^T(Ax' + Bx - f) = 0 \tag{6.1a}$$

$$Z_2^T \mathcal{F}x - Z_2^T g = 0 \tag{6.1b}$$

is an index one DAE with the same solutions as the original DAE. Thus, differentiating (6.1b) once in the stabilized sense, we obtain the completion

$$
\begin{aligned}
Z_{1,0}^T(Ax' + Bx - f) &= 0 && (6.2a)\\
(Z_2^T\mathcal{F}x - Z_2^T g)' &= (-\lambda I)(Z_2^T\mathcal{F}x - Z_2^T g) && (6.2b)
\end{aligned}
$$

Or, in a more explicit form it is

$$
\begin{aligned}
Z_{1,0}^T Ax' &= -Bx + f && (6.3a)\\
Z_2^T\mathcal{F}x' &= [-\lambda Z_2^T\mathcal{F} - (Z_2^T\mathcal{F})']x + (Z_2^T g) + (Z_2^T g)' && (6.3b)
\end{aligned}
$$

Note that the system (6.2) can be viewed as a special case of (5.34) with $C_1 = C_2 = 0$. Therefore, Theorem 9 and 10 hold for this completion. Moreover, since now $Z_{1,0}$ and Z_2 are independent of λ, D^{-1} and U do not depend on λ as well. Thus, by Remark 1, we conclude that the completion (6.2) is asymptotically stable for larger λ's. Note that λ is allowed to be time variable.

6.2 Computation Using Least Squares

As mentioned earlier, the left hand side of (6.2b) now involves the derivative of Z_2. Since Z_2 is calculated numerically, the computation of Z_2' based on Z_2 can create an unacceptable amount of numerical error. We can avoid this by computing the completion (6.2) using least squares method that does not involve Z_2'. We have already outlined the process for the constant coefficient case in Chapter 4. We will modify that process for the time variable case now.

Let

$$
R = \begin{bmatrix} Z_1^T K \\ Z_2^T K \\ Z_3^T \\ Z_4^T \end{bmatrix}
$$

where Z_1^T, Z_2^T are defined as before, $Z_3^T = (Z_2^T)'K + Z_2^T S + \lambda Z_2^T K$, and Z_4^T are extra rows to make R invertible. We will calculate the LSC obtained from

$$
R\widehat{J}w = R(-\widehat{\mathcal{F}}x + \widehat{g}) \qquad (6.4)
$$

where

$$\widehat{J} = \begin{bmatrix} A & 0 & 0 & 0 & \cdots & 0 \\ A' + B & A & 0 & 0 & \cdots & 0 \\ A'' + 2B' & 2A' + B & A & 0 & \cdots & 0 \\ A''' + 3B'' & 3A'' + 3B' & 3A' + B & A & \cdots & 0 \\ \vdots & & & & \ddots & \vdots \end{bmatrix}, \widehat{\mathcal{F}} = \begin{bmatrix} B \\ B' \\ B'' \\ \vdots \\ B^k \end{bmatrix}, \widehat{g} = \begin{bmatrix} f \\ f' \\ f'' \\ \vdots \\ f^k \end{bmatrix}$$

This is a special case of (5.5) with $\mathcal{D} = \frac{d}{dt}$. Note that the LSC of (6.4) is given by the first block row of

$$(R\widehat{J})^{\dagger} R(-\widehat{\mathcal{F}}x + \widehat{g})$$

Considering previous calculations for $D = \frac{d}{dt}$, we obtain $[R\widehat{J}, R(-\widehat{\mathcal{F}}x + \widehat{g})]$ as

$$\left[\begin{array}{c|ccccc|c} Z_{1,0}^T A & 0 & 0 & \cdots & 0 & 0 & Z_{1,0}^T(Bx + f) \\ 0 & 0 & 0 & \cdots & 0 & 0 & Z_2^T(\mathcal{F}x + g) \\ -Z_2^T\mathcal{F} & 0 & 0 & \cdots & 0 & 0 & (Z_2^T\mathcal{F})' + \lambda(Z_2^T\mathcal{F}) + (Z_2^T g)' + \lambda(Z_2^T g) \\ \hline 0 & R_1 & R_2 & \cdots & \cdots & \cdots & \cdots \end{array}\right] \tag{6.5}$$

Note that since $\begin{bmatrix} Z_{1,0}^T A \\ -Z_2^T\mathcal{F} \end{bmatrix} = \begin{bmatrix} Z_1^T K \\ Z_3^T \end{bmatrix} \tilde{J}V^T$, and is invertible, and since $Z_2^T K(\tilde{J}V^T) = 0$, the rows of $Z_1^T K, Z_2^T K, Z_3^T$ are linearly independent. Also, since $\tilde{J}V^T$ has n columns, we have

$$n = \text{rank}(\begin{bmatrix} Z_1^T K \\ Z_2^T K \\ Z_3^T \end{bmatrix} \tilde{J}V^T) \leq \text{rank}(\tilde{J}V^T) \leq n$$

Therefore,

$$\text{corank}(\tilde{J}V^T) = (k+2)n - \text{rank}(\tilde{J}V^T) = (k+1)n = \dim(Z_2^T K) + \dim(Z_4^T)$$

This implies that we can choose Z_4 such that both R is invertible and $Z_4^T(\tilde{J}V^T) = R_0 = 0$. In this special circumstance we have from [24] that

$$\begin{bmatrix} X & 0 \\ Y & Z \end{bmatrix}^{\dagger} = \begin{bmatrix} X^{\dagger} & 0 \\ -X^{\dagger}YZ^{\dagger} & Z^{\dagger} \end{bmatrix}$$

so that the Moore-Penrose inverse of the block lower triangular matrix is also block lower triangular matrix. Thus the first block row gives the LSC

$$x' = \begin{bmatrix} Z_{1,0}^T A \\ -Z_2^T \mathcal{F} \end{bmatrix}^{-1} \left(\begin{bmatrix} Z_{1,0}^T B \\ (\lambda I)(Z_2^T \mathcal{F}) + (Z_2^T \mathcal{F})' \end{bmatrix} x + \begin{bmatrix} Z_{1,0}^T f \\ (\lambda I)(Z_2^T g) + (Z_2^T f)' \end{bmatrix} \right) \tag{6.6}$$

which is equivalent to

$$Z_{1,0}^T (Ax' + Bx - f) = 0 \tag{6.7}$$

$$(Z_2^T \mathcal{F} x + Z_2^T g)' = (-\lambda I)(Z_2^T \mathcal{F} x + Z_2^T g) \tag{6.8}$$

which is the same as (6.2). Note that the matrix R does not involve Z_2'.

6.3 The Smoothness of Calculations

While we have proved the existence of Z_1, Z_2, we also need to show that they can be obtained smoothly. In the constant coefficient case Z_1, Z_2 are constant, therefore smoothness was not an issue. However, they can be time dependent in the time varying case since the derivative array is time variable. We will demonstrate a process to obtain Z_1, Z_2 smoothly. Since the solutions of ordinary differential equations are smooth, we will construct an ODE system that includes Z_1 and Z_2 as particular solutions. We will illustrate the process only for Z_2 here. It can be done similarly for Z_1 as well.

Given a smooth jacobian $J(t)$ with constant rank, suppose that we want to find a smooth Z_2 as in the hypothesis. For a fixed $t_0 \in I$, by the singular decomposition theorem, there exist constant unitary matrices U_0 and V_0 such that

$$U_0^T J(t_0) V_0 = \begin{bmatrix} \Sigma_0 & 0 \\ 0 & 0 \end{bmatrix} \tag{6.9}$$

where Σ_0 is nonsingular. Then, from [42], the constant matrices U_0 and V_0 can be extended to smooth matrix functions $U(t) = \begin{bmatrix} \widehat{Z}(t) & Z(t) \end{bmatrix}$ and $V(t) = \begin{bmatrix} \widehat{T}(t) & T(t) \end{bmatrix}$ such that

$$\begin{bmatrix} \widehat{Z}(t) & Z(t) \end{bmatrix}^T J(t) \begin{bmatrix} \widehat{T}(t) & T(t) \end{bmatrix} = \begin{bmatrix} \Sigma(t) & 0 \\ 0 & 0 \end{bmatrix} \tag{6.10}$$

with $\begin{bmatrix} \widehat{Z}(t_0) & Z(t_0) \end{bmatrix} = U_0$, $\begin{bmatrix} \widehat{T}(t_0) & T(t_0) \end{bmatrix} = V_0$. Moreover, the columns of $\widehat{Z}(t)$ and $Z(t)$ span the range and corange of J(t) respectively. In other words, the matrix function $Z(t)$ satisfies the hypothesis of $Z_2(t)$ we are trying to calculate.

Now consider the following ordinary differential equation

$$\begin{bmatrix} \widehat{Z}(t)^T J(t) \\ T(t)^T \end{bmatrix} T'(t) = -\begin{bmatrix} \widehat{Z}(t)^T J'(t) T(t) \\ 0 \end{bmatrix} \tag{6.11a}$$

$$\begin{bmatrix} \widehat{T}(t)^T J(t)^T \\ Z(t)^T \end{bmatrix} Z'(t) = -\begin{bmatrix} \widehat{T}(t)^T J'(t)^T Z(t) \\ 0 \end{bmatrix} \tag{6.11b}$$

$$\begin{bmatrix} T(t)^T \\ \widehat{T}(t)^T \end{bmatrix} \widehat{T}'(t) = -\begin{bmatrix} T'(t)^T \widehat{T}(t) \\ 0 \end{bmatrix} \tag{6.11c}$$

$$\begin{bmatrix} Z(t)^T \\ \widehat{Z}(t)^T \end{bmatrix} \widehat{Z}'(t) = -\begin{bmatrix} Z'(t)^T \widehat{Z}(t) \\ 0 \end{bmatrix} \tag{6.11d}$$

with initial conditions $\begin{bmatrix} \widehat{Z}(t_0) & Z(t_0) \end{bmatrix} = U_0$ and $\begin{bmatrix} \widehat{T}(t_0) & T(t_0) \end{bmatrix} = V_0$. A straightforward calculation shows that the matrices in (6.10) satisfy this ODE and the initial conditions [42]. Thus, by the uniqueness of the solutions, $Z_2^T(t) = Z(t)$ can be calculated as part of the solution of this ODE.

Chapter 7

Conclusions and Future Research

7.1 Conclusion

Completing a DAE to an ODE has been a major approach for numerically solving DAEs for twenty years. This is especially advantageous for unstructured higher index systems since the direct application of numerical methods to DAEs require the problem to have lower index and special structure. There are many ways one can obtain a completion given a DAE. The basic idea of EI is to numerically compute a completion by solving derivative array equations using least squares methods. While the EI approach has several advantages when it comes to efficient implementation, the behavior of the additional dynamics can be a problem [23]. Ideally, the solution manifold will be asymptotically stable within the open set defined by the solutions of the completion. In other words, we would like the additional solutions to tend towards the manifold.

Our goal in this thesis has been to analyze the additional dynamics of LSCs for linear DAEs and modify the completion process to have better dynamics. We can outline the analytical completion process as follows

- Form the derivative array equations using the DAE
- Form the least squares equations using the derivative array
- Solve the least squares equations to obtain an analytical formula for the dynamical part of the completion.

We first analyzed the additional dynamics of LSC defined by the standard derivative array. By the standard derivative array we mean the derivative array obtained by successively differentiating the DAE. We started our investigation with linear constant coefficient systems. Using a canonical decomposition, we first identified the part of the DAE that creates the additional dynamics. We then applied the completion process outlined above and obtained an analytical formula for the dynamical part of the DAE using linear algebraic techniques. For the constant coefficient case, the dynamics are determined by the eigenvalues of the system. We have proved that the eigenvalues that control the additional dynamics are all equal to "zero". This means that the additional solutions will move away from the solution manifold at a polynomial speed whose degree is equal to the index of the nilpotent matrix in the canonical decomposition of the DAE. While this is not the worst case possible, it is not what we are looking for either. The situation is more complex for time varying systems. We demonstrated through an example that the additional dynamics can be arbitrary for the LTV systems if one uses the standard derivative array.

One way to change the behavior of additional dynamics is to modify the derivative array prior to applying the least squares method. One candidate was to use stabilized differentiation. Instead of simply differentiating the DAE, we connect each new derivative with all the previous equations using a parameter λ. We have proved that the additional eigenvalues of such a completion are then given by $-\lambda$. This means that the additional solutions will converge to the solution manifold for any λ positive. Thus the stability can be ensured in advance. However, the eigenvalues $-\lambda$ has a Jordan block of size equal to the index of the DAE. We then applied the method to time varying systems. In the time varying case λ is allowed to be time variable. We have showed that the solution manifold is stable if λ satisfies a certain inequality that involves A, B and some other coefficients. We have also analyzed the convergence of solutions to a specific solution of the DAE and showed that this might require even more restrictions on λ.

Another technique we have developed to improve the additional dynamics is the alternative stabilized completion. Given a linear DAE, we first reformulate the system as a semi-explicit index-1 DAE using auxiliary matrices Z_1, Z_2. We then differentiate the constraint equation in the stabilized sense to obtain a completion. It is then a straightforward task to analyze the stability of the completion. For the constant coefficient case, the manifold is asymptotically stable for any λ positive and there is no jordan block accompanying the eigenvalues $-\lambda$. In other words, the additional dynamics are pure exponential. While a positive λ is not sufficient for the stability in the LTV case, we still have the stability for a every sufficiently large λ, which is not the case with the first technique. One difficulty with the alternative stabilized completion in the LTV case is the presence of Z_2' in the completion and smooth calculation of Z_1 and Z_2, for which we have outlined techniques to overcome the problem. We use the least squares method to avoid Z_2' and use an ODE to compute Z_1, Z_2 smoothly.

In summary, we have developed two methods to alter the additional dynamics of LSCs in the desired direction. While each method has its own analytical and numerical advantages, we believe that they will at least jointly provide the necessary tools to obtain similar results for nonlinear systems.

Chapter 8

Future Research

We have analyzed the additional dynamics of LSC for linear DAEs and developed two methods to produce LSCs with desired stability properties. Ultimately our goal is two extend these results to nonlinear DAEs. It is obvious that nonlinear analysis will contain additional difficulties. However, having obtained results for linear time varying systems is promising in that respect since LTV systems have many structural similarities with nonlinear DAEs. One way to start would be constructing linear jacobians by using linearization techniques. One can then obtain a stabilized LTV completion and investigate the relationship to the original nonlinear DAE. The nonlinearity will probably make the results more local in nature.

In this thesis we only concerned ourselves with the effect of the dynamics and took only the coefficients into the consideration in the analysis of the stability. We did not examine the effect of constant terms. While in the long term the eigenvalues are what determines the stability, the constant term also becomes important if we were to consider the behavior of dynamics in a finite interval. An important question is then how λ should be chosen to achieve a certain approximation in a given interval. This will depend not only on the coefficient matrices but also on the function $f(t)$ and the interval.

Our work in this thesis has been fully theoretical in nature. There are also a number of computational issues to be examined. One of them is technical issues in the selection process for the stability parameter λ [2]. While we specified in the thesis what assumptions λ has to satisfy, the results concern only the limit case, therefore an ideal selection of λ might require a technical analysis based on the structure of the specific problem, area of integration and performance goals. Also, Theorem 10 was proved using Gronwall and some related inequalities, which are known to often give very conservative results. A future research topic is to prove versions of Theorem 10 under weaker, more practical assumptions on λ. One other numerical issue is to calculate the components of the alternative completion smoothly, for which we outlined a process. There are also general numerical issues regarding the computational cost of LSC process in general, such as calculation of consistent initial conditions, derivative arrays and solving nonlinear equations that are formed by Newton's iteration.

Although we have constructed stabilized completions only for linear DAEs, it is also possible to apply the same techniques to many other systems that have different forms. Consider the following DAE system with delay

$$Ax' + Bx + Cx(t - \tau) + f(t) = 0 \tag{8.1}$$

or

$$Ax' + Bx + Cx(t - \tau) + f(\tau) = 0$$

Assuming that (A, B) is regular, we can apply the techniques we have developed to obtain a stabilized completion of the form

$$x' = \Theta x + \Sigma \tag{8.2}$$

where Σ contains $f, x(t - \tau)$ and their derivatives. One can then investigate the conditions necessary or modify the completion so that the higher derivatives of $x(t - \tau)$ will disappear so (8.2) will be an ODE delay system whose analysis is much easier than (8.1). The stability of the completion will enable (8.2) to effectively approximate (8.1).

Another possible example is the control systems with observers. Suppose that we have a control DAE

$$Ax' + Hx = Bu \tag{8.3}$$

where A is singular, and would like to obtain an observer $y = Cx + Du$. Provided that (A, B) is regular, we can obtain again a stabilized completion of (8.3) of the form

$$\tilde{x}' + \tilde{H}\tilde{x} = Bv \tag{8.4}$$

It is easier to find an observer for this ordinary control system. Let

$$\tilde{y} = \tilde{C}\tilde{x} + \tilde{D}u \tag{8.5}$$

be an observer for (8.4). Then, since (8.4) is a stabilized completion we have

$$\|\tilde{x} - x\| \to 0$$

which implies

$$\|\tilde{y} - y\| \to 0$$

thus (8.5) can be used as an observer for (8.3). In fact, since the additional dynamics have to converge to the DAE dynamics, we only need the solution manifold of (8.3) be observable and not all of the solutions of (8.4) be observable.

We should also note that while we have considered only continuous systems in the thesis, the techniques can be modified for a discreet system

$$Ax_{n+1} + Bx_n = f_n$$

with A being singular. There are some similarities as well as differences between continuous time and discreet time analysis. For example, if A and B are constant, using the iterative scheme we can produce the equations

$$Ax_{n+1} + Bx_n = f_n \tag{8.6}$$

$$Ax_{n+2} + Bx_{n+1} = f_{n+1} \tag{8.7}$$

$$\vdots = \vdots \tag{8.8}$$

$$Ax_{n+k} + Bx_{n+k-1} = f_{n+2} \tag{8.9}$$

which gives the system

$$Jw = \mathcal{F}x_n + g_n \tag{8.10}$$

analogous to a derivative array, where $w = [x_{n+1}, x_{n+1}, \cdots, x_{n+k}]$. Note that J will have the same structure as in the continuous time case when A, B are constant. Therefore, using similar assumptions we can solve the systems for x_{n+1} in terms of x_n and g, which will be a discrete system of the form

$$x_{n+1} + Cx_n = g_n \tag{8.11}$$

whose analysis is much easier.

If A and B are time variable, then the Jacobian in (8.10) will have a different but simpler structure than its continuous counterpart since we will not have terms coming form the product rule. Note that every new differentiation produces additional solutions. However, an iteration creates only an equivalent system.

Chapter 9

Publications and Presentations

Publications

- I. Okay, S. L. Campbell, and P. Kunkel, *The Additional Dynamics of Least Squares Completions for Linear Differential Algebraic Equations*, Linear Algebra and Its Applications, 425 (2007) 471-485.

- I. Okay, S. L. Campbell, and P. Kunkel, *Completions of Implicitly Defined Vector Fields and Their Applications*, Proc. MTNS 2008, To appear.

- I. Okay, S. L. Campbell, and P. Kunkel, *Stabilized Least Squares Completions for Linear Time Varying Differential Algebraic Equations*, (In Preparation).

Presentations

- *Fifth International Conference on Dynamic Systems and Applications*
  May 2007, Morehouse College Atlanta, Georgia, USA

- *Southeastern-Atlantic Regional Conference on Differential Equations*
  October 2007, Murray State University Murray, Kentucky, USA

- *Mathematical Theory of Network and Systems*
  July 2008, Virginia Tech Blacksburg, Virginia

Bibliography

[1] C. Arévalo, S. L. Campbell, and M. Selva Soto, *Unitary Partitioning in General Constraint Preserving DAE Integrators*, Mathematical and Computer Modelling, 40 (2004), 1273-1284.

[2] U. Ascher, H. Chin and S. Reich, *Stabilization of DAEs and invariant manifolds* , Numer. Math. 67 (1994), 131-149.

[3] Uri M. Ascher, L. R. Petzold, *Computer Methods for Ordinary Differential Equations and Differential Algebraic Equations*, SIAM, (1998).

[4] A. Barrlund, *Constrained least squares methods for linear time varying DAE systems*, Numer. Math. 60 (1991), 145-161.

[5] A. Barrlund, *Constrained least squares methods for nonlinear DAE systems*, UMINF-91.15, Institute of Information Processing, University of Umeøa, Sweden (1991).

[6] A. Barrlund, B. Kågström, *Analytical and numerical solutions to higher index linear variable coefficient DAE systems*, J. Comp. Appl. Math.

[7] J. Baumgarte, *Stabilization of constraints and integrals of motion in dynamical systems*, Comp. Math. Appl. Mech. Eng. 1(1972), 1-16.

[8] A. Ben-Israel, *A Newton-Raphson method for the solution of systems of equations*, J. Math. Anal. Appl. 15 (1966), 243-252.

[9] K. E. Brenan, S. L. Campbell, and L. R. Petzold, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, SIAM, 1996.

[10] S. L. Campbell, *Singular Systems of Differential Equations*, Pitman, 1980.

[11] S. L. Campbell, *Singular Systems of Differential Equations*, Pitman, 1982.

[12] S. L. Campbell, *One canonical form for higher index linear time varying singular systems*, Circuits System Signal Process., 2 (1983), 311-326.

[13] S. L. Campbell, *The numerical solution of higher index linear time varying singular systems of differential equations*, SIAM J. Sci. Stat. Comp., 6 (1985), 334-348.

[14] S. L. Campbell, *Rank deficient least squares and the numerical solution of linear singular implicit systems of differential equations*, in Linear Algebra and Its Role in Systems Theory, Contemporary Mathematics, 47 (1985), 51-63.

[15] S. L. Campbell, *Consistent initial conditions for linear time varying singular systems*, in Frequency Domain and State Space Methods, Edited by C. I. Byrnes and A. Lindquist, Elsevier, (1986), 313-318.

[16] S. L. Campbell, *A general form for solvable linear time varying singular systems of differential equations*, SIAM J. Math. Anal., 18 (1987), 1101-1115.

[17] S. L. Campbell, *A computational method for general higher index singular systems of differential equations*, 1989 IMACS Transactions Scientific Computing, Vol. 1.2 (1989), 555-560.

[18] S. L. Campbell, *Least squares completions for nonlinear index three Hessenberg DAEs*, Computational and Applied Mathematics, II Differential Equations, Edited by W. F. Ames and P. J. van der Houwen, North-Holland, 1992, 117-126.

[19] S. L. Campbell, *Uniqueness of completions for linear time varying differential algebraic equations*, Linear Algebra & Its Appl., 161 (1992), 55-68.

[20] S. L. Campbell, *Least squares completions for nonlinear differential algebraic equations*, Numer. Math., 65 (1993), 77-94.

[21] S. L. Campbell and C. W. Gear, *The index of general nonlinear DAEs*, Numerische Mathematik, 72 (1995), 173-196.

[22] S. L. Campbell and E. Griepentrog, *Solvability of general differential algebraic equations*, SIAM J. Scientific Computation, 16 (1995), 257-270.

[23] S. L. Campbell and B. Leimkuhler, *Differentiation of constraints in differential algebraic equations*, Mechanics of Machines and Structures, 19 (1991), 19-39.

[24] S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*, Dover, 1991.

[25] S. L. Campbell, E. Moore, *Progress on a numerical method for nonlinear higher index DAEs II*, Circuit Systems and Sugnal Processing.

[26] S. L. Campbell and E. Moore, *Constraint preserving integrators for general nonlinear higher index DAEs*, Numerische Mathematik, 69 (1995), 383-399.

[27] S. L. Campbell, E. Moore, Y. Zhong, *Utilization of automatic differentiation in control algorithms*, Proc. 31st IEEE Conf. Decision Control.

[28] S. L. Campbell, E. Moore, and Y. Zhong, *Constraint preserving integrators for unstructured higher index DAEs*, Zeit. fuer Ang. Math. und Mech. (ZAMM), 76 (1996) S1, 83-86.

[29] S. L. Campbell, N. K. Nichols, W. J. Terrell, *Duality, observality and controllability for linear time varying descriptor systems*, Circuit Systems and Sugnal Processing, 10 (1991), 455-470.

[30] S. L. Campbell, L. R. Petzold, *Canonical forms and solvable singular systems of differential equations*, SIAM J. Algebraic Discrete Methods, 4 (1983), 517-521.

[31] S. L. Campbell, W. J. Terrell, *Observality of linear time varying descriptor system*, SIAM J. Matrix Anal., 12 (1991), 484-496.

[32] S. L. Campbell and K. D. Yeomans, *Behavior of the nonunique terms in general DAE integrators*, Appl. Numer. Math., 28 (1998), 209-226.

[33] S. L. Campbell and Y. Zhong, *Jacobian reuse in explicit integrators for higher index DAEs*, Applied Numerical Mathematics, 25 (1997), 391-412.

[34] K. D. Clark, *A Structural form for higher index semistate equations I:Theory and applications to circuit and control*, Linear Algebra Appl., 98 (1988), 169-197.

[35] I. Duff, C. W. Gear, *Computing the structural index*, SIAM J. Alg. Disc. Methods, 7(1986), 594-603.

[36] F. R. Gantmatcher, *The Theory of Matrices*, Volume II, Chelsea, 1964.

[37] C. W. Gear, *Maintaining solution invariants in the numerical solutions of ODEs*, SIAM J. Sci. Stat. Comput. 7(3) (1986), 734-743.

[38] C. W. Gear, *Differential algebraic equations index transformation*, SIAM J. Sci. Stat. Comput. 7 (1988) 39-47.

[39] C. W. Gear, B. Leimkuhler, G. K. Gupta, *Automatic integration of Euler-Lagrange equations with constraints*, J. Comput. Appl. Math., 12/13 (1985), 77-90.

[40] E. Griepentrog, *Index reduction methods for differential algebraic equations*, Seminarberichte Nr. 92-1., Humboldt-Universität zu Berlin Fachbereich Mathematik, (1992), 14-29.

[41] E. Hairer, C. Lubich, and M. Roche, *The numerical solution of differential-algebraic systems by Runge-Kutta methods*, Springer Lecture Notes in Mathematics No. 1409, 1989.

[42] P. Kunkel, V. Mehrmann, *Differential-Algebraic Equations: Analysis and Numerical Solutions*, EMS Textbooks in Mathematics, Europian Mathematical Society, Zurich 2006.

[43] P. Kunkel, V. Mehrmann, and W. Rath, *A new software package for linear differential algebraic equations*, SIAM J. Sci. Comp., to appear.

[44] P. Kunkel and V. Mehrmann, *Canonical forms for linear differential-algebraic equations with variable coefficients*, Journal of Computational and Applied Mathematics, Vol. 56, 1994, pp. 225-251.

[45] P. Kunkel and V. Mehrmann, *A new class of discretization methods for the solution of linear differential llgebraic equations with variable coefficients*, SIAM Journal Numerical Analysis, Vol. 33, No.5, October 1996, pp. 1941-1961.

[46] P. Kunkel, V. Mehrmann, W. Rath and J. Weickert, *GELDA: A software package for the solution of general linear differential algebraic equations*, SIAM Journal Scientific Computing, Vol. 18, 1997, pp. 115 - 138.

[47] D. G. Luenberger, *Dynamic equations in descriptor form*, IEEE Trans. Automat. Control, AC-22 (1977), 312-321.

[48] S. Mattsson, G. Söderlind, *Index reduction in differential algebraic equations using dummy derivatives*, SIAM J. Sci. Statist. Comput., 14 (1993), 677-692.

[49] C. C. Pantelides, *The consistent initilization of differential algebraic systems*, SIAM J. Sci. Statist. Comput., 9 (1988), 213-231.

[50] L. R. Petzold, C. W. Gear, *ODE methods for the solutions of differential algebraic systems*, Sandia Report, 82-8051, 1982.

[51] F. A. Potra and W. C. Rheinboldt, *Differential-geometric techniques for solving differential algebraic equations*, in *Real-Time Integration Methods for Mechanical System Simulation* Eds. E. J. Haug and R. C. Deyo, Springer-Verlag Computer and Systems Sciences Vol. 69, 1991, 155-191.

[52] P.J. Rabier, W.C Rheinboldt, *A general existence and uniqueness theorem for implicit differential algebraic equations*, Diff. Int. Eqns. 4, (1991), 563-582.

[53] P.J. Rabier, W.C Rheinboldt, *A geometric treatment of implicit differential algebraic equations*, J. Diff. Eqns. 109, (1994), 110-146.

[54] S. Reich, *On a geometric interpretation of differential-algebraic equations*, Circ. Sys. Sig. Process, 9 (1990), 367-382.

[55] S. Reich, *On a differential geometric characterization of differential algebraic equations Part II: Reduction Methods*, Tech report 09-03-90, Technische University Dresden, Sektion Informationstechnik, (1990).

[56] S. Reich, *On an existence and uniqueness theory for nonlinear differential-algebraic equations*, Circuits, Systems Signal Process, 10 (1991), 343-359.

[57] W. Rheinboldt, *Differential-algebraic systems as differential equations on manifolds*, Math. Comp. 43 (1984), 473-482.

[58] W. Rheinboldt, *On the existence and uniqueness of solutions of nonlinear semi-explicit differential algebraic equations*, Nonlinear Anal., Theory, Methods, Appl., 16 (1991), 647-661.

[59] L. M. Silverman, *Inversion of multivariable systems*, IEEE Trans. Automat. Control, AC-14 (1969), 270-276.

[60] J. H. Wilkinson, *Note on the practical significance of the Drazin inverse*, in *Recent Applications of Generalized Inverses*, edited by S. L. Campbell, Pitman, 1982.