

# Theory and Application of a Class of Abstract Differential-Algebraic Equations

Mark A. Pierson

Dissertation submitted to the Faculty of the  
Virginia Polytechnic Institute and State University  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy  
in  
Mathematics

Terry L. Herdman, co-Chair  
Eugene M. Cliff, co-Chair  
Jeffrey T. Borggaard  
John A. Burns  
David L. Russell

April 25, 2005  
Blacksburg, Virginia

Keywords: abstract differential-algebraic equations (DAE), partial differential-algebraic equations (PDAE), hybrid systems, systems of partial differential equations, existence and uniqueness, well-posedness

Copyright 2005, Mark A. Pierson

# Theory and Application of a Class of Abstract Differential-Algebraic Equations

Mark A. Pierson

(ABSTRACT)

We first provide a detailed background of a geometric projection methodology developed by Professor Roswitha März at Humboldt University in Berlin for showing uniqueness and existence of solutions for ordinary differential-algebraic equations (DAEs). Because of the geometric and operator-theoretic aspects of this particular method, it can be extended to the case of infinite-dimensional abstract DAEs. For example, partial differential equations (PDEs) are often formulated as abstract Cauchy or evolution problems which we label *abstract ordinary differential equations* or AODE. Using this abstract formulation, existence and uniqueness of the Cauchy problem has been studied. Similarly, we look at an AODE system with operator constraint equations to formulate an *abstract differential-algebraic equation* or ADAE problem. Existence and uniqueness of solutions is shown under certain conditions on the operators for both index-1 and index-2 abstract DAEs. These existence and uniqueness results are then applied to some index-1 DAEs in the area of thermodynamic modeling of a chemical vapor deposition reactor and to a structural dynamics problem. The application for the structural dynamics problem, in particular, provides a detailed construction of the model and development of the DAE framework. Existence and uniqueness are primarily demonstrated using a semigroup approach. Finally, an exploration of some issues which arise from discretizing the abstract DAE are discussed.

This research was funded in part by the Defense Advanced Research Projects Agency (DARPA), the National Aeronautical and Space Administration (NASA) Langley Research Center (LaRC), and the National Institute of Aerospace (NIA).

# Dedication

This work is dedicated to Professor Terry Herdman who first got me interested in mathematics at the Virginia Tech Northern Virginia Falls Church Campus and was gracious enough to allow my further pursuits in Blacksburg. I would also dedicate this to my wife Rebekah Paulson for giving me the time and support to follow my dream. Finally, I dedicate this work to my father, Richard E. Pierson, Lt. Col., U.S.A.F. (ret.), who is a Ph.D. at heart and is pursuing his own research dreams.

# Acknowledgments

I especially want to thank Professor Gene Cliff who, since he is a Professor Emeritus in the Department of Aerospace and Ocean Engineering, was able to spend plenty of time trying to keep me honest about the mathematics and in keeping me on track to reach my goal. I also would like to acknowledge the many excellent professors that I had the pleasure of taking classes from who taught me a lot about mathematics: Professors Terry Herdman, David Russell, John Burns, Joe Ball, Jeff Borggaard and Peter Haskell. Obviously, any mistakes I have made in this work are not their fault but solely my own.

I also want to acknowledge my previous superiors at the Office of Naval Research (ONR), Drs. Spiro Lekoudis and Ron DeMarco, who motivated me to get back into the academic arena through stimulating intellectual discussions. I also want to thank them for allowing me time off to do my own research through the Research Opportunities for Program Officers available at ONR and for the necessary time to take late afternoon classes at the Virginia Tech Northern Virginia Campus.

Finally, I want to thank Professor Terry Herdman again for all he has done to help me at Virginia Tech both as a student and for his advice in the job search process. Despite times of discouragement, he still had confidence in me that I would persevere and continue to make progress. The fact that you never gave up on me helped keep me going.

# Contents

- 1 Introduction** **1**
  - 1.1 The ‘What’ and ‘Why’ of DAEs . . . . . 1
  - 1.2 Matrix Inverses . . . . . 4
  - 1.3 A Geometric Approach . . . . . 6
  - 1.4 Abstract Differential Equations . . . . . 7
  - 1.5 Partial Differential-Algebraic Equations (PDAEs) and Index . . . . . 8
  - 1.6 Discretization Issues . . . . . 9
  
- 2 Ordinary DAEs: The Finite-dimensional LTV Case** **10**
  - 2.1 Background and Matrix Sequence . . . . . 10
  - 2.2 Key Results . . . . . 12
  
- 3 Abstract DAEs: Infinite-dimensional LTI Index-1 Case** **22**
  - 3.1 Infinite-dimensional Linear Algebra . . . . . 22
  - 3.2 Abstract Index and Operator Sequence . . . . . 23
  - 3.3 The Index-1 Semi-explicit Case . . . . . 24
  - 3.4 A Case Where  $\mathcal{B}$  is Bounded . . . . . 27
  - 3.5 A Case Where  $\mathcal{B}$  is Unbounded . . . . . 34
  
- 4 Abstract DAEs: Infinite-dimensional LTI Index-2 Case** **38**
  - 4.1 Background for Abstract Hessenberg DAE . . . . . 38

4.2	A Case Where $\mathcal{B}$ is Bounded . . . . .	43
4.3	A Case Where $\mathcal{B}$ is Unbounded . . . . .	46
<b>5</b>	<b>Applications</b>	<b>49</b>
5.1	Thermal Model for Chemical Vapor Deposition Reactor . . . . .	49
5.2	Coupled Transversal Motion of Two Beams Connected by a Rigid Joint . . . . .	51
5.2.1	Modeling and Development of Abstract DAE . . . . .	52
5.2.2	Existence and Uniqueness . . . . .	61
<b>6</b>	<b>Discretized Differential-Algebraic Equations</b>	<b>70</b>
6.1	Solvability of the Differential-Algebraic Equation . . . . .	70
6.2	Incorporating the Algebraic Constraints into an Explicit ODE . . . . .	74
6.3	Hessenberg Form of DAEs . . . . .	77
<b>7</b>	<b>Conclusions</b>	<b>81</b>
7.1	Summary and Contributions . . . . .	81
7.2	Future Work . . . . .	82
	<b>Bibliography</b>	<b>84</b>
	<b>Vita</b>	<b>87</b>

# List of Figures

5.1	Simplified Thermal Model of a Chemical Vapor Deposition Reactor . . .	50
5.2	An Example of a Truss Array . . . . .	52
5.3	Model of Two Beams and a Joint . . . . .	53

# Chapter 1

## Introduction

### 1.1 The ‘What’ and ‘Why’ of DAEs

Differential-algebraic equations can now be found quite frequently in the literature . A myriad of terminology is often used depending on the application in which they arose. For example, one will often run across terms such as descriptor systems, singular differential equations, degenerate differential equations, semi-state systems, implicit differential equations, constrained differential equations and generalized state space systems. They all refer to the same concept of differential-algebraic equations. So what are they?

Essentially differential-algebraic equations can be thought of as a differential equation describing a process that is coupled with an algebraic constraint equation. Hence, its name. Differential-algebraic equations, hereafter referred to as DAEs, of this sort take the form

$$\begin{aligned} F(x'(t), x(t), y(t), t) &= 0 \\ G(x(t), y(t), t) &= 0, \end{aligned}$$

where the Jacobian  $F_{x'}$  is nonsingular. Note also that one of the variables  $y(t)$  is not differentiated in a differential equation and is coupled to the differential equation  $F(\cdot, \cdot, \cdot, \cdot) = 0$  via the algebraic constraint equation  $G(\cdot, \cdot, \cdot) = 0$ . This nondifferentiated variable is often a key identifier that one may have a DAE. More generally, one could have a fully implicit DAE of the form

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

where in this case the Jacobian  $F_{y'}$  is singular. This is one reason for the name singular differential equations. The name singular differential equations or degenerate differential equations also comes about from the following form which we will primarily use in this



paper:

$$E(t)x'(t) + B(t)x(t) = q(t),$$

where  $x(t)$  and  $q(t)$  are finite-dimensional vectors and the finite-dimensional matrix  $E(t)$  is singular for all  $t$  of interest. This is the linear time-varying (LTV) form of a DAE. We also have the case of linear constant coefficients or the linear time-invariant (LTI) form

$$Ex'(t) + Bx(t) = q(t),$$

where the matrix  $E$  is singular and thus noninvertible. This is the primary form we will focus our research on in the abstract Hilbert space setting. We will then further refine the LTI DAE form into a block matrix equation which is called the semi-explicit form (since the  $x'(t)$  variable is available explicitly with no singular coefficient):

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x'_1(t) \\ x'_2(t) \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} q_1(t) \\ q_2(t) \end{bmatrix}$$

where  $E = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$  is singular.

However, this results in two equations: a differential equation and an algebraic constraint equation.

$$\begin{aligned} x'(t) + B_{11}x_1(t) + B_{12}x_2(t) &= q_1(t) \\ B_{21}x_1(t) + B_{22}x_2(t) &= q_2(t), \end{aligned}$$

where the variable  $x_2(t)$  is not differentiated. The ability to be able to do such a *decoupling* for a general DAE will be important later on in determining existence and uniqueness of solutions to the DAE. We will especially use this property in the more abstract setting in Chapters 3 and 4.

It is also important to note that in some cases one can solve for the  $x_2(t)$  variable in the algebraic constraint equation explicitly in terms of  $x_1(t)$  (such as when  $B_{22}$  is invertible, for example). This is then substituted back into the differential equation to obtain an ordinary differential equation (ODE) in terms of only  $x_1(t)$  without any constraints. This is often referred to as the “underlying” differential equation. However, one needs to be careful in the sense that while solutions of the original DAE are solutions to the underlying ODE assuming the same initial conditions, the opposite is not true for all initial conditions. All solutions of the ODE are not necessarily solutions to the DAE. In particular, there is no guaranteed existence of solutions under similar hypotheses as for ODEs. This raises the issue of what is called *consistent initial conditions*. In other words, the initial conditions need to be compatible with the algebraic constraint equation. Furthermore, when one solves the underlying ODE, even using consistent initial conditions, there tends to be the problem of numerical “drift”. In particular the numerical

solution drifts off of the manifold determined by the algebraic constraint equation. Thus, the final numerical solution may no longer satisfy the algebraic constraints. Moreover, DAEs are numerically similar to very stiff ODEs and hence are not always numerically solvable. This issue of a DAE not being an ODE is addressed in a frequently cited paper by Petzold [35] with the appropriate title, *Differential/Algebraic Equations are not ODEs*.

Now that we have answered the basic question of what is a DAE? Why do we need to be concerned with solving them? Why and where do they come up? First of all, as discussed above, we may not be able to convert every DAE to its underlying ODE. Even if we were able to accomplish this it may be too expensive or time consuming to do so. Even if we could easily convert it to an ODE, we may lose whatever sparsity or structure that we initially may have had in the original DAE formulation. Often the variables have some physical significance which then gets lost in the translation to an ODE. It then gets hard to interpret the results of the underlying ODE in terms of the original variables. We have a similar situation in the case where a parametrization is related to the original variables but then the useful relationship gets lost in the transformation. It also turns out that the dimension of the solution manifold is different for the DAE versus the underlying ODE. How does this impact the results? Finally, with the advent of automatic computer generation of differential equations and algebraic constraint equations by various software packages, it would be advantageous to have an automated solver for DAEs that can then be coupled together with these other off-the-shelf packages. Hence, it is necessary that a methodology be developed to solve DAEs directly without using the associate underlying ODE.

DAEs arise frequently in numerous applications. In fact, in certain applications, DAEs provide a better framework for modeling and analysis. DAEs can be found in a rich variety of applications such as: constrained variational problems, chemical reaction kinetics and processes, combustion, electrical circuits and networks, trajectory prescribed path control, optimal control problems, robotics, discretization of partial differential equations (PDEs) using the method of lines, and mechanical systems simulations, to name just a few. Applications also arise in the area of PDEs such as from incompressible Navier-Stokes equations when using finite element methods for spatial discretization where the algebraic constraints arise from the divergence-free condition. We will apply some of our theory of abstract DAEs to some example applications involving PDEs later in Chapter 5.

There are many good monographs and books that can give one a good introduction to both the theory and the numerical solution of DAEs. A short list would include works by Brenan, Campbell and Petzold [3]; Campbell [5] and [6]; Ascher and Petzold [1]; Boyarintsev [4]; Rabier and Rheinboldt [36]; Hairer, Norsett and Wanner [20]; and März [33].

## 1.2 Matrix Inverses

Of significance is the fact that most of the methods used to discuss either the theory or the numerical solution of DAEs involve the definition of some sort of inverse for matrices or operators. The conventional definition of an *inverse* of a matrix  $A$ , i.e.,  $A^{-1}$ , applies to the case of a nonsingular square matrix where  $A^{-1}A = AA^{-1} = I$ . For a nonsingular square matrix  $A$ , its inverse is unique. If we have a singular square matrix  $A$ , then we have the concept of the *Drazin inverse*  $A^D$  which must satisfy the following conditions:

$$\begin{aligned} AA^D &= A^D A \\ A^D AA^D &= A^D \\ (I - A^D A)A^k &= 0, \end{aligned}$$

where  $k$  is the index of the matrix  $A$ , that is,  $k$  is the smallest nonnegative integer satisfying the relation  $\text{rank } A^{k+1} = \text{rank } A^k$ . If  $k = 0$ , then the matrix  $A$  is nonsingular and  $A^D = A^{-1}$ . The Drazin inverse  $A^D$  is unique for any given square matrix  $A$ . The following representation for the Drazin inverse may be useful. First put the matrix  $A$  into its Jordan canonical form:

$$A = N \begin{bmatrix} J_0 & 0 \\ 0 & J_1 \end{bmatrix} N^{-1},$$

where  $J_0$  consists of nilpotent blocks and  $J_1$  consists of nonsingular blocks. Every matrix  $A$  can be put into its Jordan canonical form (see Gantmacher [16] for example). The Drazin inverse  $A^D$  can be represented by (see Boyarintsev [4])

$$A^D = N \begin{bmatrix} 0 & 0 \\ 0 & J_1^{-1} \end{bmatrix} N^{-1}.$$

What if a matrix is not square? For rectangular matrices, we have the Moore-Penrose *pseudo-inverse*,  $A^+$ , where  $A^+ = X$  satisfies the four Moore-Penrose conditions:

$$\begin{aligned} AXA &= A \\ XAX &= X \\ (AX)^T &= AX \\ (XA)^T &= XA. \end{aligned}$$

The Moore-Penrose inverse  $A^+$  is also unique for any given matrix  $A$ . Thus, in the case where  $A$  is a square nonsingular matrix,  $A^{-1} = A^+$ . If a rectangular matrix  $A$  has full column rank, then  $A^+ = (A^T A)^{-1} A^T$ . If it has full row rank instead, we have  $A^+ = A^T (A A^T)^{-1}$ . The pseudo-inverse  $A^+$  also has the properties  $(A^+)^+ = A$  and

$(A^T)^+ = (A^+)^T$ . If the matrix  $A$  is real, symmetric and square, i.e.,  $A^T = A$ , then the pseudo-inverse  $A^+$  is also symmetric. Furthermore, for real, symmetric square  $A$ , we have that the Drazin inverse  $A^D = A^+$  the pseudoinverse of  $A$ . Finally, the Moore-Penrose conditions essentially result in the requirement that  $AA^+$  and  $A^+A$  become orthogonal projections onto  $\text{im } A$  and  $\text{im } A^T$  respectively (see Golub and Van Loan [18]). The use of projection operators will appear frequently throughout our development of the theory of DAEs in abstract spaces.

We can generalize the concept of the inverse of a matrix  $A$  even further. These more general concepts of an inverse become useful for the infinite-dimensional operator setting where the notion of square and rectangular do not exactly fit. The most general form of an inverse is called a *semi-inverse* by Boyarintsev [4]. The semi-inverse for any  $m \times n$  matrix  $A$  is defined as the  $n \times m$  matrix  $A^\sim$  such that for  $A^\sim = X$  we have  $AXA = A$ . However, one drawback is that the semi-inverse  $A^\sim$  of a matrix  $A$  is generally not unique. A representation for the semi-inverse where  $A^\sim = X$  is

$$X = +(I - A^\sim A)U + V(I - AA^\sim),$$

where  $A^\sim$  is any semi-inverse matrix for  $A$  and  $U$  and  $V$  are arbitrary matrices of the appropriate dimension. Thus, if you know one semi-inverse of  $A$ , all others can be found via the representation above. We now become a little more restrictive and define what Boyarintsev calls the *inverse semi-reciprocal* matrix  $A^-$  of an arbitrary  $m \times n$  matrix  $A$ . The inverse semi-reciprocal  $A^-$  satisfies both  $AA^-A = A$  and  $A^-AA^- = A^-$ . März [33] calls this the *reflexive generalized inverse* of the matrix  $A$ . We will follow März and use reflexive generalized inverse as this name seems more appropriate. Of course, both terms are translated into English from their native languages. The reflexive generalized inverse  $A^-$  is also generally not unique. It has a representation of the more complicated form

$$X = A^-AA^- + A^-AU(I - AA^-) + (I - A^-A)UAA^- + (I - A^-A)UAU(I - AA^-),$$

where  $A^-$  is any reflexive generalized inverse of  $A$  and  $U$  is an arbitrary matrix of the appropriate dimension. Thus, if you know one reflexive generalized inverse of  $A$ , all others can be found via the representation above. Note also since we have added one more condition  $XAX = X$ , we have one less arbitrary matrix, i.e., just  $U$  vice both  $U$  and  $V$ , in the representation formula. We also remark that if  $A$  is square and has index  $k = 1$ , then the Drazin inverse  $A^D$  is also a reflexive generalized inverse of  $A$ . It should be clear that both the semi-inverse  $A^\sim$  and the reflexive generalized inverse  $A^-$  exist for any arbitrary matrix whether rectangular or square, invertible or noninvertible. As indicated these inverses are not unique.

As we will see, the concept of a semi-inverse or reflexive generalized inverse becomes very useful for defining various subspaces of  $m \times n$  rectangular matrices. These reflexive generalized inverses can also be readily extended to operators on infinite-dimensional

spaces. We will use some of the subspace descriptions which follow in the case of finite-dimensional matrices and for bounded infinite-dimensional operators. For example, Boyarintsev [4] provides the following results. We have  $\ker A = \text{im}(I - A\tilde{A}) = \ker A\tilde{A}$  and  $\text{im} A = \text{im} AA\tilde{A}$ . If we have two matrices, an  $m_1 \times n$  matrix  $A$  and an  $m_2 \times n$  matrix  $B$ , then  $\ker A \cap \ker B = \text{im} P$  where  $P$  is a projector on  $\mathbb{R}^n$  defined by  $P = (I - A\tilde{A})[I - (B - BA\tilde{A})(B - BA\tilde{A})\tilde{A}]$ . Likewise,  $\ker A \cup \ker B = \ker \tilde{P}$  where  $\tilde{P} = [I - (B - BA\tilde{A})(B - BA\tilde{A})\tilde{A}]B$ . Finally, for  $m \times n_1$  matrix  $A$  and  $m \times n_2$  matrix  $B$ , we have  $\text{im} A \cup \text{im} B = \ker Q$  where  $Q = [I - (B - AA\tilde{B})(B - AA\tilde{B})\tilde{A}](I - AA\tilde{A})$ .

### 1.3 A Geometric Approach

The subspace descriptions above become more meaningful in light of using a geometrical approach to solving DAEs. Rabier and Rheinboldt [36] use a geometric approach for rigid mechanical systems. The geometric framework turns out to be very powerful. While their approach is very useful in many finite-dimensional applications it does not seem to extend as well to the infinite-dimensional framework. This then becomes the advantage of the geometric approach utilized by März [33]. She defines operators, projectors and subspaces derived from the matrices  $E(t)$  and  $B(t)$  which are found in the DAE  $E(t)x'(t) + B(t)x(t) = q(t)$ . Furthermore, at the end of her article [33], she hints at a framework for applying her methodology to the infinite-dimensional setting. It is her methodology that we will follow in this research. We will also utilize the same notation as that in her article for the operators, projectors and subspaces in order to provide for comparison and to avoid confusion in translation to another set of symbols.

Chapter 2 below repeats the key Lemmas and Theorem from [33] for the LTV index-2 DAE. We also will define there the concept of an index using März's definitions. We further include the proofs of these Lemmas and Theorem in Chapter 2 as some of the needed proofs of the Propositions and Lemmas are not in her paper but are instead in other hard to find references. In addition, those proofs that are contained in [33] often leave out many steps due to page limit constraints and it is not always obvious how to fill in the missing steps. It was considered important to fill in the gaps to provide a more complete picture of this methodology before proceeding to the abstract case. Thus, there is no new work represented by Chapter 2 other than that of putting it all together in a unified treatment. We will then build from there starting with the infinite-dimensional setting beginning in Chapter 3.

We will see that one of the primary benefits of März's geometric approach is that this method can decouple the DAE into its *inherent regular ODE* and the algebraic constraint equation. In the case of an index-2 DAE which is more complicated we can even determine the so-called "hidden constraint equation." This inherent regular ODE is usually located

in a lower dimensional manifold than that of the “underlying ODE” discussed earlier. Hence, the inherent regular ODE is also referred to as the *essential underlying ODE* by other authors such as Petzold because it incorporates only the essential or smallest part of the space that is needed. In the infinite-dimensional setting, this inherent regular ODE takes the form of an inhomogeneous abstract Cauchy problem (IACP) for the LTI case, i.e.,  $x'(t) + \mathcal{A}x(t) = f(t)$  or  $x'(t) = -\mathcal{A}x(t) + f(t)$ . In the case where the operator  $\mathcal{A}(t)$  depends on time as well we have evolution equations. We use the term abstract ODE or AODE to describe either situation.

## 1.4 Abstract Differential Equations

In order to determine whether a unique solution exists to the AODE we end up using the already developed theory of differential equations in abstract spaces. This theory is usually developed in the most general abstract infinite-dimensional Banach space setting. However, since we are just beginning to develop a theory we will restrict ourselves to the Hilbert space setting. Obviously any result which applies to a Banach space, or a reflexive Banach space, will also apply to the Hilbert space setting. Many good monographs and books exist on the general theory of abstract differential equations such as: Krein [25]; Ladas and Lakshmikantham [27]; Zaidman [41]; and Fattorini [13] and [14]. One of the methods that will be utilized to show existence and uniqueness of the AODE in this research includes that of semigroup theory. The primary monographs and books used here for semigroup theory and applications include: Pazy [34]; Goldstein [17]; Kato [23]; Engel and Nagel [12]; Liu and Zheng [30]; and the original main reference in this area, Hille and Phillips [21]. Engel and Nagel was particularly useful since it is an up-to-date treatment of semigroups in a textbook format and includes many of the advances in the theory since the other classic monographs were written.

There is also some literature on *degenerate* or *singular* Cauchy problems. These include the monographs by Carroll and Showalter [8] and Favini and Yagi [15]. However, these problems differ considerably from those which will be discussed here. Our abstract DAE takes the form  $\mathcal{E}(t)x'(t) + \mathcal{B}x(t) = q(t)$  where the operator  $\mathcal{E}(t)$  is singular for all  $t$  and the prime indicates differentiation with respect to time  $t$ . Carroll and Showalter look at abstract equations of the form  $\mathcal{A}(t)x''(t) + \mathcal{B}(t)x'(t) + \mathcal{C}(t)x(t) = g(t)$  where some of the operators  $\mathcal{A}(t)$ ,  $\mathcal{B}(t)$  and/or  $\mathcal{C}(t)$  become zero or infinite at  $t = 0$ . They define the problem as *degenerate* if some of the operator coefficients become zero as  $t \rightarrow 0$  and as *singular* if at least one of the operator coefficients becomes infinity as  $t \rightarrow 0$ . This distinction becomes somewhat artificial in the case when some of the operators are invertible or when a suitable change of variable is introduced. Their study was motivated by many problems in physics, geometry, and applied mathematics that are of this form. Similarly, Favini and Yagi were also motivated by the large number of partial differential

equations arising in physics and the applied sciences which can be written in this form. They studied similar problems of the form  $\frac{d}{dt}(\mathcal{E}x) = \mathcal{L}x + f(t)$  where  $\mathcal{E}^{-1}$  may exist but is not necessarily bounded. In our case, the distinction is that for an abstract DAE  $\mathcal{E}$  is noninvertible for all time  $t$ , not just at  $t = 0$  or when the inverse is unbounded.

In addition to various inverses of matrices, another aspect of DAEs that has been studied is that of *matrix pencils*. For a LTI DAE of the form  $Ex'(t) + Bx(t) = q(t)$  one can look at the matrix pencil  $(\lambda E + B)$ . It turns out that the DAE is solvable if and only if the matrix pencil is a regular pencil, i.e.,  $\det(\lambda E + B)$  is not identically zero as a function of  $\lambda$ . Gantmacher [16] discusses matrix pencils in general and develops what he calls the *Kronecker canonical form*. Brenan, Campbell and Petzold [3] and Campbell [5] discuss solvability of LTI DAEs with respect to matrix pencils. This concept can be extended as well to the abstract differential equation case. This area was started with Keldysh [24] and was further developed by Markus [32] for a spectral theory of *polynomial operator pencils*. Yakubov and Yakubov [39] extended the idea further to that of *generalized resolvents*. The idea is to invert the polynomial operator pencil in order to derive the solution of the abstract differential equation. Yakubov and Yakubov, in particular, use this methodology to investigate higher order elliptic, parabolic and hyperbolic differential equations.

The closest discussion of abstract differential equations related to abstract DAEs is that of Zaidman [42]. In this monograph, Zaidman investigates abstract differential equations of the form  $\frac{d}{dt}(\mathcal{E}x(t)) = \mathcal{B}x(t)$  or  $\mathcal{E}x'(t) = \mathcal{B}x(t)$  where  $\mathcal{E}$  and  $\mathcal{A}$  are linear, often unbounded, operators on either a Hilbert space or a Banach space. He calls these type of equations *singular abstract differential equations*. As above, in most cases it is assumed that the operator  $\mathcal{E}$  is invertible (but not necessarily bounded) and that the operator  $\mathcal{E}^{-1}\mathcal{B}$  exists. In another case, it is determined that even if  $\mathcal{E}$  is not invertible, that the operator pencil  $(\lambda\mathcal{E} + \mathcal{B})$  is invertible in some half-plane  $Re\lambda > \omega$ . Finally, all cases are for the homogeneous form of the equation and do not address the inhomogeneous case. It would be interesting in the future to compare Zaidman's two methodologies with that of März. Zaidman does not address the issue of index and it may be that the different approaches are required because of different index equations.

## 1.5 Partial Differential-Algebraic Equations (PDAEs) and Index

There is also now a lot of literature in the area of what is called *partial differential-algebraic equations*. This term is used to describe various DAEs that arise from PDEs. For example, Lucht, Strehmel and Eichler-Liebenow [31] label those systems where Fourier or Laplace transforms of PDEs result in a sequence of DAEs as PDAEs. They give

an example from a model of population dynamics. PDAEs usually take the form of  $Au_t(t, y) + Bu_{yy}(t, y) + Cu(t, y) = q(t, y)$  where at least one of the matrices  $A, B, C \in L(\mathbb{R}^n)$  is singular. Lucht, et al, further indicate that there are two kinds of indexes for PDAEs: a differential spatial index and a uniform differential time index. März's group defines a PDAE as a set of PDEs coupled with ordinary DAEs. März's definition of abstract index generalizes the Kronecker index. Essentially the notion of abstract DAE index by März approximates that of the uniform differential time index. We will not investigate the various definitions of index for PDAEs and abstract DAEs. We will use that given by März as it proves useful in determining existence and uniqueness of solutions. A detailed analysis of index for PDAEs can be found in Campbell and Marszalek [7]. Finally, some work has also been done by Reid, Lin and Wittkopf [37] in the area of existence and uniqueness of PDAEs using differential elimination - completion algorithms. They show that for autonomous first-order DAEs, this algorithm is equivalent to the Cartan-Kuranishi algorithm for completing a system of differential equations to involutive form. This algorithm can also be applied to PDAEs and results in existence and uniqueness for systems in involutive form.

The framework developed by März [33] and by Lamour, März and Tischendorf [29] for abstract DAEs covers the case of PDAEs. However, it is much more general in that it covers other cases as well including those often considered to be *hybrid systems*. We therefore emphasize this methodology in our research toward existence and uniqueness theorems for abstract DAEs. Chapters 3 and 4 builds up the theory for abstract DAEs. The examples provided in Chapter 5 indicate the diversity of problems that can utilize this approach.

## 1.6 Discretization Issues

Finally, in Chapter 6 we briefly look at issues related to numerical solutions of abstract DAEs. When discretizing the spatial variables using finite element methods, we end up with an ordinary DAE. We look at the Hessenberg form in particular and discuss various formulations such as the underlying ODE, the inherent regular ODE, and differentiation of constraints to reduce the index of the resulting DAEs. We then end with some final remarks, conclusions and future directions for further research.



# Chapter 2

## Ordinary DAEs: The Finite-dimensional LTV Case

### 2.1 Background and Matrix Sequence

Consider the linear time-varying DAE:

$$A(t)(D(t)x(t))' + B(t)x(t) = q(t), \quad t \in \mathcal{I}, \quad (2.1)$$

with matrix coefficients

$$A \in C(\mathcal{I}, L(\mathbb{R}^n, \mathbb{R}^m)), \quad D \in C(\mathcal{I}, L(\mathbb{R}^m, \mathbb{R}^n)), \quad B \in C(\mathcal{I}, L(\mathbb{R}^m)),$$

where the prime notation in (2.1) represents differentiation with respect to time  $t$ .

**Definition 2.1. Subspace and Matrix Sequence** (We use the notation of März [33])

We define the following matrices and subspaces which will be used in this section:

$$\begin{aligned} G_0(t) &:= A(t)D(t), \quad B_0(t) = B(t) \\ \text{For } i = 0, 1 : \\ N_i(t) &:= \ker G_i(t), \\ Q_i(t) &\in L(\mathbb{R}^m), \quad Q_i(t)^2 = Q_i(t), \quad \text{im } Q_i(t) = N_i(t), \\ P_i(t) &:= I - Q_i(t), \\ W_i(t) &\in L(\mathbb{R}^m), \quad W_i(t)^2 = W_i(t), \quad \ker W_i(t) = \text{im } G_i(t), \\ S_i(t) &= \ker W_i(t)B_i(t) = \{z \in \mathbb{R}^m : B_i(t)z \in \text{im } G_i(t)\}, \\ G_{i+1}(t) &:= G_i(t) + B_i(t)Q_i(t), \quad B_{i+1}(t) := B_i(t)P_i(t). \end{aligned}$$

**Definition 2.2.** From März [33], we define the *reflexive generalized inverse* of  $D(t)$  by  $D(t)^-$  such that  $DD^-D = D$  and  $D^-DD^- = D^-$ . More information on generalized inverses can be found in Boyarintsev [4] where he refers to matrices such as  $D(t)^-$  as the *inverse semi-reciprocal matrix*. In general,  $D(t)^-$  exists for any  $m \times n$  matrix and is not unique. However, we can make it unique by setting  $D(t)^-D(t) = P_0(t)$  with  $P_0(t) = I - Q_0(t)$ . We then define the unique projector  $R$  by  $R(t) = D(t)D(t)^-$  using this unique  $D(t)^-$ .

*Remark.* There is more than one choice for  $Q_0(t)$  that will satisfy the conditions in Definition 2.1. However, once a particular  $Q_0(t)$  is chosen,  $D(t)^-$  is fixed and unique. Thus,  $D(t)^-$  will depend on the choice of  $Q_0(t)$  just as  $P_0(t)$  does.

**Definition 2.3.** The ordered pair of continuous matrix functions  $A$  and  $D$  is said to be *well-matched* if

$$\text{im } D(t) \oplus \ker A(t) = \mathbb{R}^n, t \in \mathcal{I}, \quad (2.2)$$

and these subspaces are spanned by continuously differentiable bases.

When  $A(t)$  and  $D(t)$  are well-matched, there is a unique projector  $R(t) \in L(\mathbb{R}^n)$  such that,

$$R^2(t) = R(t), \quad \text{im } R(t) = \text{im } D(t), \quad \ker R(t) = \ker A(t), \quad R \in C^1(\mathcal{I}, L(\mathbb{R}^n)).$$

The decomposition (2.2) implies that for all  $t \in \mathcal{I}$ ,

$$\text{im } A(t)D(t) = \text{im } A(t), \quad \ker A(t)D(t) = \ker D(t), \quad \ker A(t) \cap \text{im } D(t) = \{0\}.$$

Note that both matrices  $A(t)$  and  $D(t)$  will have constant rank  $r$  on  $\mathcal{I}$ .

**Definition 2.4.** The DAE (2.1) with well-matched  $A(t)$  and  $D(t)$  and nontrivial  $N_0(t)$  has *index-1* if  $N_0(t) \cap S_0(t) = \{0\}$  on  $t \in \mathcal{I}$  and has *index-2* if  $\dim(N_0(t) \cap S_0(t)) = \text{constant} > 0$  and  $N_1(t) \cap S_1(t) = \{0\}$  on  $t \in \mathcal{I}$ . The intersection  $N_j(t) \cap S_j(t)$  will always have the same dimension as the subspace  $N_{j+1}(t)$ ,  $j = \{0, 1\}$ .

In terms of  $G_i(t)$ , the DAE (2.1) has *index-1* if  $G_1(t)$  is nonsingular and has *index-2* if  $G_1(t)$  is singular with constant rank  $< m$  and  $G_2(t)$  is nonsingular.

Before we proceed any further we will need the following proposition derived from Griepentrog and März [19] (Lemma 14, Appendix A). This proposition is established using the same steps as the proof of the Lemma in [19], but it used in a different setting with different results. We provide the proof of the proposition here as it is not found anywhere else in the published literature.

**Proposition 2.1.** *Given an index-2 DAE (2.1), let  $Q_1$  be any projection onto  $N_1$ . Then  $\tilde{Q}_1 = Q_1G_2^{-1}BP_0$  is the projection onto  $N_1$  along  $S_1$ .*

*Proof.* Using the subspace and matrix sequence from Definition 2.1 we have  $N_1 = \ker G_1$  and  $N_1 = \text{im } Q_1$ . Thus,  $G_1 Q_1 = 0$ . Since  $Q_1$  is a projection,  $Q_1^2 = Q_1$ . Additionally,  $G_2 := G_1 + B_1 Q_1$  and since the DAE has index-2,  $G_2$  is invertible. Then,

$$\begin{aligned}
I - Q_1 &= G_2^{-1} G_2 (I - Q_1) \\
&= G_2^{-1} (G_1 + B_1 Q_1) (I - Q_1) \\
&= G_2^{-1} G_1 (I - Q_1) + G_2^{-1} B_1 Q_1 (I - Q_1) \\
&= G_2^{-1} G_1 (I - Q_1) \\
&= G_2^{-1} G_1 - G_2^{-1} (G_1 Q_1) \\
&= G_2^{-1} G_1.
\end{aligned}$$

Thus,  $I - Q_1 = G_2^{-1} G_1$ . We next have

$$\begin{aligned}
Q_1 &= I - (I - Q_1) = I - G_2^{-1} G_1 \\
&= I - G_2^{-1} (G_2 - B_1 Q_1) = I - I + G_2^{-1} B_1 Q_1 \\
&= G_2^{-1} B_1 Q_1
\end{aligned}$$

or,  $Q_1 = G_2^{-1} B_1 Q_1$ . Now it follows that

$$\begin{aligned}
(Q_1 G_2^{-1} B_1)^2 &= Q_1 (G_2^{-1} B_1 Q_1) G_2^{-1} B_1 \\
&= Q_1^2 G_2^{-1} B_1 = Q_1 G_2^{-1} B_1.
\end{aligned}$$

This identity shows that  $Q_1 G_2^{-1} B_1$  is a projector with  $\text{im } Q_1 G_2^{-1} B_1 \subseteq \text{im } Q_1 = N_1$ . Then  $Q_1 G_2^{-1} B_1 x = 0$  implies  $G_2^{-1} B_1 x \in \text{im } (I - Q_1)$  or  $G_2^{-1} B_1 x = (I - Q_1)z$  for some  $z$ . Thus,

$$\begin{aligned}
B_1 x &= G_2 (I - Q_1) z = (G_1 + B_1 Q_1) (I - Q_1) z \\
&= G_1 (I - Q_1) z + B_1 Q_1 (I - Q_1) z \\
&= G_1 z \in \text{im } G_1.
\end{aligned}$$

This implies  $x \in S_1$  since  $S_1 = \{x : B_1 x \in \text{im } G_1\}$ . Since  $\ker Q_1 G_2^{-1} B_1 = S_1$  and  $\text{im } Q_1 G_2^{-1} B_1 \oplus \ker Q_1 G_2^{-1} B_1 = \mathbb{R}^m$ , we have  $\text{im } Q_1 G_2^{-1} B_1 = \ker G_1 = N_1$  (Note that  $N_1 \oplus S_1 = \mathbb{R}^m$ ). Finally, since  $B_1 := B P_0$ , we end up with  $\tilde{Q}_1 = Q_1 G_2^{-1} B P_0$  where  $\tilde{Q}_1$  is the projection onto  $N_1$  along  $S_1$  starting with a  $Q_1$  being any projection onto  $N_1$ .  $\square$

## 2.2 Key Results

We now state and prove some Lemmas and a main Theorem which come from März [33]. The proof of Lemma 1 is expanded from that of Balla and März [2]. This reference is also not readily available so the proof is provided here for completeness.

**Lemma 2.1.** *Given an index-2 DAE (2.1). Let  $Q_1(t)$  denote the projection onto  $N_1(t)$  along  $S_1(t)$ ,  $t \in \mathcal{I}$ . Then, for all  $t \in \mathcal{I}$ , we have the decomposition*

$$D(t)S_1(t) \oplus D(t)N_1(t) \oplus \ker A(t) = \mathbb{R}^n, \quad (2.3)$$

where

$$D(t)P_1(t)D(t)^-, \quad D(t)Q_1(t)D(t)^-, \quad I - R(t)$$

are the uniquely determined projectors respectively that realize this decomposition.

If additionally,  $D(t)S_1(t)$  and  $D(t)N_1(t)$  are spanned by continuously differentiable functions defined on  $\mathcal{I}$ , then  $DP_1D^-, DQ_1D^- \in C^1(\mathcal{I}, L(\mathbb{R}^n))$ .

*Proof.* The projectors  $\widehat{P}_1(t)$  and  $\widehat{Q}_1(t)$  are determined from the matrix sequence in Definition 2.1. Let  $\widehat{Q}_1(t)$  be any projector onto  $N_1(t)$ . From Proposition 2.1 it is shown that if we start with any non-unique projector  $\widehat{Q}_1(t)$  onto  $N_1(t)$  we can obtain the unique projector  $Q_1(t)$  onto  $N_1(t)$  along  $S_1(t)$  by the formula  $Q_1 = \widehat{Q}_1 G_2^{-1} B P_0$ . (We drop the function of  $t$  for ease of notation.)

By construction,  $Q_1 = Q_1 G_2^{-1} B \widehat{P}_0$ ,  $G_2 = G_1 + B \widehat{P}_0 Q_1$ , and  $Q_1 \widehat{Q}_0 = 0$ , where  $\widehat{Q}_0$  is any projector onto  $N_0 = \ker G_0 = \ker AD$ . We get  $Q_1 \widehat{Q}_0 = 0$  because  $Q_0$  is onto  $N_0$  with  $N_0 \subseteq S_1$  while  $Q_1$  is onto  $N_1$  along  $S_1$ . Since  $P_1 = I - Q_1$ , we have

$$\widehat{P}_0 P_1 \widehat{Q}_0 = \widehat{P}_0 (I - Q_1) \widehat{Q}_0 = \widehat{P}_0 \widehat{Q}_0 - \widehat{P}_0 Q_1 \widehat{Q}_0 = 0.$$

Therefore,

$$\begin{aligned} (DQ_1D^-)(DQ_1D^-) &= DQ_1(D^-D)Q_1D^- = DQ_1\widehat{P}_0Q_1D^- \\ &= DQ_1(I - \widehat{Q}_0)Q_1D^- \\ &= DQ_1Q_1D^- - D(Q_1\widehat{Q}_0)Q_1D^- \\ &= DQ_1D^-. \end{aligned}$$

Similarly,

$$\begin{aligned} (DP_1D^-)(DP_1D^-) &= D(I - Q_1)D^-D(I - Q_1)D^- \\ &= DD^-D(I - Q_1)D^- - DQ_1D^-D(I - Q_1)D^- \\ &= D(I - Q_1)D^- - DQ_1D^-DD^- + (DQ_1D^-)(DQ_1D^-) \\ &= D(I - Q_1)D^- - DQ_1D^- + DQ_1D^- \\ &= D(I - Q_1)D^- \\ &= DP_1D^- \end{aligned}$$

and  $(I - R)(I - R) = I - R$ .

Obviously,  $(DQ_1D^-)(DP_1D^-) = 0$  and  $(DP_1D^-)(DQ_1D^-) = 0$ . Moreover,  $R = DD^-$  leads to

$$\begin{aligned} DQ_1D^-(I - R) &= DQ_1D^- - DQ_1D^-DD^- = DQ_1D^- - DQ_1D^- = 0 \quad \text{and,} \\ (I - R)DQ_1D^- &= 0. \end{aligned}$$

Similarly,  $DP_1D^-(I - R) = 0$  and  $(I - R)DP_1D^- = 0$ . Thus,  $DP_1D^-$ ,  $DQ_1D^-$ , and  $I - R$  are projectors in  $\mathbb{R}^n$ .

Since  $P_1$  and  $Q_1$  are orthogonal projectors and  $Q_1$  projects onto the  $\ker G_1$ ,  $P_1$  projects onto  $\text{im } G_1$ . Furthermore, for index-2 DAEs we have the decomposition  $N_1(t) \oplus S_1(t) = \mathbb{R}^m$  and  $N_1 = \text{im } Q_1$ . Thus, it follows that  $\text{im } DP_1 \subseteq DS_1$  and  $\text{im } DQ_1 \subseteq DN_1$ .

As noted  $N_1 = \text{im } Q_1$  and  $Q_1$  projects onto the  $\ker G_1$ . Thus, if  $G_1$  has rank  $r_1$ , the subspace  $N_1$  has dimension  $m - r_1$ . In addition, since

$$P_0Q_1 = (I - Q_0)Q_1 = (I - Q_0Q_1)Q_1,$$

where  $(I - Q_0Q_1)$  is nonsingular,  $P_0N_1$  has the same dimension  $m - r_1$  as  $N_1$ . Hence,  $\dim DN_1 = \dim P_0N_1 = \dim N_1 = m - r_1$ . Therefore, we conclude that

$$\text{im } DQ_1D^- = DN_1, \quad \text{im } DP_1D^- = DS_1.$$

For the last part of the Lemma we assume both subspaces  $DS_1$  and  $DN_1$  are spanned by continuously differentiable functions. Additionally, by construction we chose  $Q_1(t)$  and  $P_1(t)$  to be continuous. We now show that the projectors  $DP_1D^-$  and  $DQ_1D^-$  are continuous.

We let  $\dim(\text{im } D(t)) = r$ , then since  $A(t)$  and  $D(t)$  are well-matched  $\dim(\ker A(t)) = n - r$ . Since  $\dim DN_1 = m - r_1$ , this implies  $\dim DS_1 = r - (m - r_1) = r + r_1 - m$ .

Now denote the continuously differentiable functions that span  $DS_1$  by  $DS_1 = \text{span}\{Ds_j : j = 1, \dots, r - (m - r_1)\}$ ,  $DN_1 = \text{span}\{Dn_j : j = 1, \dots, m - r_1\}$ , and  $\ker A = \text{span}\{\eta_{r+1}, \dots, \eta_n\}$ . Then, the matrix function

$$\Gamma := (Ds_1, \dots, Ds_{r-(m-r_1)}, Dn_1, \dots, Dn_{m-r_1}, \eta_{r+1}, \dots, \eta_n)$$

is nonsingular and belongs to  $C^1$ . We observe that

$$\Gamma I_{DS} \Gamma^{-1}, \quad \Gamma I_{DN} \Gamma^{-1}, \quad \Gamma I_A \Gamma^{-1}, \quad (2.4)$$

with block diagonal projector matrices defined as  $I_{DS} = \text{diag}(I_{r-(m-r_1)}, 0_{m-r_1}, 0_{n-r})$ ,  $I_{DN} = \text{diag}(0_{r-(m-r_1)}, I_{m-r_1}, 0_{n-r})$  and  $I_A = \text{diag}(0_{r-(m-r_1)}, 0_{m-r_1}, I_{n-r})$ , are  $C^1$  and we have the decomposition (2.3). The indices show the indicated dimensions of the identity and zero matrices. But, the projectors (2.4) are also uniquely defined by the decomposition. Therefore, from our earlier results, they coincide with  $DP_1D^-$ ,  $DQ_1D^-$ , and  $I - R$  respectively. In particular,  $DP_1D^-$  and  $DQ_1D^-$  belong to  $C^1$ .  $\square$

*Remark.* The subspace  $DS_1$  will later be relevant for the “inherent regular ODE”, whereas  $DN_1$  is related to the “hidden constraint” that appears in index-2 DAEs.

Our next goal is to develop a solution representation for the DAE (2.1) and to define the “inherent regular ODE” associated with the DAE. However, before we can go further we will need to create some identities to be used in our derivations.

From Proposition 2.1, we saw  $G_1Q_1 = 0$  and using  $Q_1 = Q_1G_2^{-1}BP_0$  we get  $Q_1Q_0 = 0$ . Therefore,

$$\begin{aligned} P_1P_0 &= G_2^{-1}G_2P_1P_0 = G_2^{-1}(G_1 + B_1Q_1)P_1P_0 \\ &= G_2^{-1}G_1P_1P_0 + G_2^{-1}B_1Q_1P_1P_0 = G_2^{-1}G_1P_1P_0 \\ &= G_2^{-1}G_1(I - Q_1)P_0 = G_2^{-1}G_1P_0 - G_2^{-1}(G_1Q_1)P_0 \\ &= G_2^{-1}G_1P_0 = G_2^{-1}(AD + BQ_0)P_0 = G_2^{-1}ADP_0 \\ &= G_2^{-1}ADD^-D = G_2^{-1}AD. \end{aligned}$$

Thus,  $G_2^{-1}AD = P_1P_0$ .

Since  $G_1 = G_0 + BQ_0 = AD + BQ_0$  by Definition 2.1 and from Proposition 2.1  $G_2^{-1}G_1 = I - Q_1$ ,

$$\begin{aligned} G_2^{-1}BQ_0 &= G_2^{-1}(G_1 - AD) = G_2^{-1}G_1 - G_2^{-1}AD \\ &= (I - Q_1) - P_1P_0 = P_1 - P_1(I - Q_0) \\ &= P_1 - P_1 + P_1Q_0 = (I - Q_1)Q_0 = Q_0 - Q_1Q_0 = Q_0. \end{aligned}$$

Thus,  $G_2^{-1}BQ_0 = Q_0$ .

Using the identity  $G_2^{-1}BP_0Q_1 = Q_1$  from Proposition 2.1 and the last identity above,

$$\begin{aligned} G_2^{-1}B &= G_2^{-1}BI = G_2^{-1}B(P_0 + Q_0) = G_2^{-1}BP_0I + G_2^{-1}BQ_0 \\ &= G_2^{-1}BP_0(P_1 + Q_1) + G_2^{-1}BQ_0 \\ &= G_2^{-1}BP_0P_1 + G_2^{-1}BP_0Q_1 + G_2^{-1}BQ_0 \\ &= G_2^{-1}BP_0P_1 + Q_1 + Q_0. \end{aligned}$$

Thus,  $G_2^{-1}B = G_2^{-1}BP_0P_1 + Q_1 + Q_0$ .

We now multiply the DAE (2.1) by  $G_2^{-1}$  and use the identity  $A = AR = ADD^-$  to get

$$\begin{aligned} G_2^{-1}A(Dx)' + G_2^{-1}Bx &= G_2^{-1}q \\ G_2^{-1}ADD^-(Dx)' + G_2^{-1}Bx &= G_2^{-1}q. \end{aligned}$$

Substituting in the identities we just derived for  $G_2^{-1}AD$  and  $G_2^{-1}B$  along with  $P_0 = D^-D$  leads to

$$\begin{aligned} P_1P_0D^-(Dx)' + G_2^{-1}BP_0P_1x + Q_1x + Q_0x &= G_2^{-1}q \quad \text{or,} \\ P_1P_0D^-(Dx)' + G_2^{-1}BD^-DP_1x + Q_1x + Q_0x &= G_2^{-1}q. \end{aligned} \tag{2.5}$$

We now multiply the modified DAE (2.5) by  $DP_1$ ,  $DQ_1$  and  $Q_0P_1$  to decouple the system into three new equations. First multiplying (2.5) by  $DP_1$  yields,

$$\begin{aligned} DP_1P_1P_0D^-(Dx)' + DP_1G_2^{-1}BD^-DP_1x + DP_1Q_1x + DP_1Q_0x &= DP_1G_2^{-1}q \\ DP_1D^-DD^-(Dx)' + DP_1G_2^{-1}BD^-DP_1x &= DP_1G_2^{-1}q \\ DP_1D^-(Dx)' + DP_1G_2^{-1}BD^-DP_1x &= DP_1G_2^{-1}q. \end{aligned} \quad (2.6)$$

where we used  $DP_1Q_0 = D(I-Q_1)Q_0 = DQ_0 - D(Q_1Q_0) = DQ_0 = 0$  since  $\text{im } Q_0 \subseteq \ker D$ .

We next multiply (2.5) by  $DQ_1$  to obtain

$$\begin{aligned} DQ_1P_1P_0D^-(Dx)' + D(Q_1G_2^{-1}BD^-D)P_1x + DQ_1^2x + D(Q_1Q_0)x &= DQ_1G_2^{-1}q \\ DQ_1P_1x + DQ_1x &= DQ_1G_2^{-1}q \\ DQ_1x &= DQ_1G_2^{-1}q, \end{aligned} \quad (2.7)$$

where the identity  $Q_1G_2^{-1}BP_0 = Q_1G_2^{-1}BD^-D = Q_1$  was used.

Finally, multiplying (2.5) by  $Q_0P_1$ ,

$$\begin{aligned} Q_0P_1P_1P_0D^-(Dx)' + Q_0P_1G_2^{-1}BD^-DP_1x + Q_0P_1Q_1x + Q_0P_1Q_0x &= Q_0P_1G_2^{-1}q \\ Q_0(I-Q_1)P_0D^-(Dx)' + Q_0P_1G_2^{-1}BD^-DP_1x + Q_0(I-Q_1)Q_0x &= Q_0P_1G_2^{-1}q \\ -Q_0Q_1D^-DD^-(Dx)' + Q_0P_1G_2^{-1}BD^-DP_1x + Q_0(Q_0-Q_1Q_0)x &= Q_0P_1G_2^{-1}q \\ -Q_0Q_1D^-(Dx)' + Q_0P_1G_2^{-1}BD^-DP_1x + Q_0x &= Q_0P_1G_2^{-1}q. \end{aligned} \quad (2.8)$$

With the above equations we can now derive the following solution representation with the new variable  $u := DP_1x$ :

$$\begin{aligned} x &= P_0x + Q_0x = D^-Dx + Q_0x = D^-D(P_1 + Q_1)x + Q_0x \\ &= D^-(DP_1x) + D^-(DQ_1x) + Q_0x. \end{aligned}$$

We substitute the results for  $DQ_1x$  from (2.7), for  $Q_0x$  from (2.8), and  $u = DP_1x$  into the last equation above to obtain

$$\begin{aligned} x &= D^-u + D^-DQ_1G_2^{-1}q + Q_0P_1G_2^{-1}q + Q_0Q_1D^-(Dx)' - Q_0P_1G_2^{-1}BD^-(DP_1x) \\ &= (I - Q_0P_1G_2^{-1}BD^-D)D^-u + P_0Q_1G_2^{-1}q + Q_0P_1G_2^{-1}q + Q_0Q_1D^-(Dx)' \\ &= (I - Q_0P_1G_2^{-1}BP_0)D^-u + (P_0Q_1 + Q_0P_1)G_2^{-1}q + Q_0Q_1D^-(Dx). \end{aligned}$$

We next work on the last term  $Q_0Q_1D^-(Dx)'$  from above:

$$\begin{aligned}
Q_0Q_1D^-(Dx)' &= Q_0Q_1D^-DD^-(Dx)' = Q_0Q_1D^-D(P_1 + Q_1)D^-(Dx)' \\
&= Q_0Q_1D^-DP_1D^-(Dx)' + Q_0Q_1D^-(DQ_1D^-)(Dx)' \\
&= Q_0[Q_1P_0P_1]D^-(Dx)' + Q_0Q_1D^-(DQ_1D^-)(D(P_1 + Q_1)x)' \\
&= Q_0[Q_1(I - Q_0)P_1]D^-(Dx)' + Q_0Q_1D^-(DQ_1D^-)(DP_1x)' \\
&\quad + Q_0Q_1D^-(DQ_1D^-)(DQ_1x)' \\
&= Q_0[Q_1P_1]D^-(Dx)' - Q_0[(Q_1Q_0)P_1]D^-(Dx)' \\
&\quad + Q_0Q_1D^-(DQ_1D^-)(DP_1x)' + Q_0Q_1P_0Q_1D^-(DQ_1x)' \\
&= Q_0Q_1D^-(DQ_1D^-)(DP_1x)' + Q_0Q_1(I - Q_0)Q_1D^-(DQ_1x)' \\
&= [Q_0Q_1D^-(DQ_1D^-DP_1x)' - Q_0Q_1D^-(DQ_1D^-)'(DP_1x)] \\
&\quad + Q_0Q_1D^-(DQ_1x)' - Q_0(Q_1Q_0)Q_1D^-(DQ_1x)' \\
&= Q_0Q_1D^-(D[Q_1P_0P_1]x)' - Q_0Q_1D^-(DQ_1D^-)'(DP_1x) \\
&\quad + Q_0Q_1D^-(DQ_1x)' \\
&= Q_0Q_1D^-(DQ_1x)' - Q_0Q_1D^-(DQ_1D^-)'(DP_1x). \tag{2.9}
\end{aligned}$$

Substituting (2.9) for  $Q_0Q_1D^-(Dx)'$ , (2.7) for  $DQ_1x$  and  $u = DP_1x$  back into our solution representation formula yields

$$\begin{aligned}
x &= (I - Q_0P_1G_2^{-1}BP_0)D^-u + (P_0Q_1 + Q_0P_1)G_2^{-1}q + Q_0Q_1D^-(DQ_1G_2^{-1}q)' \\
&\quad - Q_0Q_1D^-(DQ_1D^-)'u \\
&= [I - Q_0P_1G_2^{-1}BP_0 - Q_0Q_1D^-(DQ_1D^-)'D]D^-u + (P_0Q_1 + Q_0P_1)G_2^{-1}q \\
&\quad + Q_0Q_1D^-(DQ_1G_2^{-1}q)' \\
&= KD^-u + (P_0Q_1 + Q_0P_1)G_2^{-1}q + Q_0Q_1D^-(DQ_1G_2^{-1}q)', \tag{2.10}
\end{aligned}$$

where we used the identity  $Ru = DD^-u = DD^-DP_1x = DP_1x = u$  and defined  $K$  as follows

$$K := I - Q_0P_1G_2^{-1}BP_0 - Q_0Q_1D^-(DQ_1D^-)'D.$$

Furthermore, the  $u$  given in the solution representation formula (2.10) must satisfy the below “inherent regular ODE”.

To derive the regular inherent ODE we first note that  $DQ_0 = 0$  and

$$\begin{aligned}
DP_1D^-D &= DP_1P_0 = DP_1(I - Q_0) = DP_1 - DP_1Q_0 = DP_1 - D(I - Q_1)Q_0 \\
&= DP_1 - DQ_0 + D(Q_1Q_0) = DP_1.
\end{aligned}$$

Thus, we have

$$(DP_1x)' = (DP_1D^-Dx)' = (DP_1D^-)'(Dx) + DP_1D^-(Dx)'. \tag{2.11}$$



Substituting  $u = DP_1x$ , (2.6) for  $DP_1D^-(Dx)'$  and (2.7) for  $DQ_1x$  leads to

$$\begin{aligned} u' - (DP_1D^-)'D(P_1 + Q_1)x - [DP_1G_2^{-1}q - DP_1G_2^{-1}BD^-(DP_1x)] &= 0 \\ u' - (DP_1D^-)'u - (DP_1D^-)'(DQ_1x) + DP_1G_2^{-1}BD^-u &= DP_1G_2^{-1}q \\ u' - (DP_1D^-)'(u + DQ_1G_2^{-1}q) + DP_1G_2^{-1}BD^-u &= DP_1G_2^{-1}q. \end{aligned} \quad (2.12)$$

Of note, the ‘‘inherent regular ODE’’ (2.12) can be constructed directly from the coefficients  $A$ ,  $D$ ,  $B$ , and the right-hand side  $q$  without making any assumptions on the existence of a solution for the DAE (2.1).

*Remark.* In the index-1 case (2.10) and (2.12) simplify to

$$x = KD^-u + Q_0G_1^{-1}q, \quad K = I - Q_0G_1^{-1}B, \quad (2.13)$$

$$u' - R'u + DG_1^{-1}BD^-u = DG_1^{-1}q. \quad (2.14)$$

*Remark.* In the linear constant coefficient case ( $A$ ,  $D$ ,  $B$  do not vary with time), the solution representation formula (2.10) and the inherent regular ODE (2.12) simplify to the following as expected

$$x = KD^-u + (P_0Q_1 + Q_0P_1)G_2^{-1}q + Q_0Q_1G_2^{-1}q', \quad (2.15)$$

$$\begin{aligned} K &:= I - Q_0P_1G_2^{-1}BP_0, \\ u' + DP_1G_2^{-1}BD^-u &= DP_1G_2^{-1}q. \end{aligned} \quad (2.16)$$

Having defined the inherent regular ODE we can now prove the following useful lemma.

**Lemma 2.2.** *Given an index- $\mu$  DAE (2.1),  $\mu \in \{1, 2\}$ . Let  $DP_1D^-$ ,  $DQ_1D^-$  be continuously differentiable.*

(i) *Then, the subspaces  $D(t)S_1(t)$  and  $D(t)N_1(t)$  as well as the inherent regular ODE are uniquely determined by the problem data.*

(ii)  *$D(t)S_1(t)$  is a time-varying invariant subspace of the inherent regular ODE, i.e., if a solution belongs to this subspace at a certain point, it runs within this subspace all the time.*

(iii) *If  $D(t)S_1(t)$  and  $D(t)N_1(t)$  do not vary with time  $t$ , then, solving the IVP for (2.12) with the initial condition  $u(t_*) \in D(t_*)S_1(t_*)$  yields the same solution as solving this IVP for*

$$u' + DP_1G_2^{-1}BD^-u = DP_1G_2^{-1}q. \quad (2.17)$$

*Proof.* (From März [33].) Part (i): When defining the inherent regular ODE (2.12), the only flexibility is in our choice of  $\widehat{P}_0$  which will also affect the calculation of  $D^-$ . Suppose

then that we start with two different choices of  $P_0$  and  $\tilde{P}_0$  resulting in two different  $D^-$  and  $\tilde{D}^-$ . Then,  $D^-D = P_0$ ,  $\tilde{D}^-D = \tilde{P}_0$ , and  $DD^- = D\tilde{D}^- = R$ . Furthermore,

$$\begin{aligned}\tilde{G}_1 &= G_0 + B\tilde{Q}_0 = G_0 + BQ_0\tilde{Q}_0 = G_1(P_0 + \tilde{Q}_0), \\ (P_0 + \tilde{Q}_0)(\tilde{P}_0 + Q_0) &= I, \quad \tilde{N}_1 = (\tilde{P}_0 + Q_0)N_1, \quad \tilde{S}_1 = S_1, \\ \tilde{Q}_1 &:= (\tilde{P}_0 + Q_0)Q_1(P_0 + \tilde{Q}_0) = (\tilde{P}_0 + Q_0)Q_1, \quad \text{im } \tilde{Q}_1 = \tilde{N}_1, \quad \ker \tilde{Q}_1 = S_1, \\ D\tilde{Q}_1\tilde{D}^- &= DQ_1\tilde{D}^- = DQ_1\tilde{D}^-D\tilde{D}^- = DQ_1\tilde{D}^-DD^- = DQ_1D^-, \\ D\tilde{P}_1\tilde{D}^- &= D\tilde{D}^- - D\tilde{Q}_1\tilde{D}^- = R - DQ_1D^- = DP_1D^-.\end{aligned}$$

As a result, the projectors  $DP_1D^-$  and  $DQ_1D^-$  do not depend on the choice of  $P_0$ . Hence, their images,  $DS_1$  and  $DN_1$  respectively, also do not depend on  $P_0$ . Thus, the subspaces  $DS_1$  and  $DN_1$  are uniquely determined by the problem data.

As far as the inherent regular ODE (2.12) is concerned we need to look at the terms  $DP_1G_2^{-1}$  and  $DP_1G_2^{-1}BD^-$ :

$$\begin{aligned}\tilde{G}_2 &= G_2(P_0 + \tilde{Q}_0 + Q_0\tilde{P}_0Q_1), \quad (P_0 + \tilde{Q}_0 + Q_0\tilde{P}_0Q_1)(\tilde{P}_0 + Q_0 - Q_0\tilde{P}_0Q_1) = I, \\ \text{then, } D\tilde{G}_2^{-1} &= D(\tilde{P}_0 + Q_0 - Q_0\tilde{P}_0Q_1)G_2^{-1} = DG_2^{-1}, \\ D\tilde{P}_1\tilde{G}_2^{-1} &= D\tilde{P}_1\tilde{D}^-D\tilde{G}_2^{-1} = DP_1D^-DG_2^{-1} = DP_1G_2^{-1},\end{aligned}$$

$$\begin{aligned}D\tilde{P}_1\tilde{G}_2^{-1}B\tilde{D}^- &= D\tilde{P}_1\tilde{G}_2^{-1}B\tilde{D}^-D\tilde{D}^- = D\tilde{P}_1\tilde{G}_2^{-1}B\tilde{D}^-DD^- = D\tilde{P}_1\tilde{G}_2^{-1}B\tilde{P}_0D^- \\ &= D\tilde{P}_1G_2^{-1}BD^- = DP_1G_2^{-1}BD^-.\end{aligned}$$

Therefore, the inherent regular ODE is uniquely determined by the data.

Part (ii): Assume we have a solution  $\tilde{u} \in C^1(\mathcal{I}, \mathbb{R}^m)$  of the inherent regular ODE (2.12) such that  $\tilde{u}(t_0) \in D(t_0)S_1(t_0)$  for some  $t_0 \in \mathcal{I}$ . Put  $\tilde{u}$  into the inherent regular ODE (2.12) and then multiply both sides by  $(I - DP_1D^-)$  [since  $DP_1D^-$  is the projection onto  $DS_1$ ]. This results in

$$(I - DP_1D^-)\tilde{u}' - (I - DP_1D^-)(DP_1D^-)'\tilde{u} = 0.$$

Let  $\tilde{v} := (I - DP_1D^-)\tilde{u}$  and substitute into the above equation to yield

$$\tilde{v}' - (I - DP_1D^-)'\tilde{v} = 0, \quad \tilde{v}(t_0) = 0.$$

The solution of this auxiliary ODE is  $\tilde{v}(t) = 0$ . Thus, the solution  $\tilde{u}$  does not leave the  $D(t)S_1(t)$  subspace.

Part (iii): Even though  $D(t)P_1(t)D(t)^-$  and  $D(t)Q_1(t)D(t)^-$  are the unique projectors whose images are the constant subspaces  $D(t)S_1(t)$  and  $D(t)N_1(t)$ , these projectors are

not necessarily constant themselves since their kernels could be changing with respect to time. However, if the subspaces  $D(t)S_1(t)$  and  $D(t)N_1(t)$  are constant, there do exist constant projectors, say  $U$  and  $V$  respectively, that project onto these constant subspaces. This implies

$$\begin{aligned} (DP_1D^-)'U &= (DP_1D^-U)' = (U)' = 0, \\ (DQ_1D^-)'V &= (DQ_1D^-V)' = (V)' = 0 \quad \text{and,} \\ (DP_1D^-)'V &= (DP_1D^-V)' = 0, \quad (DQ_1D^-)'U = (DQ_1D^-U)' = 0. \end{aligned}$$

Furthermore, since  $u(t) = D(t)P_1(t)x(t)$ , we have  $u(t) = Uu(t)$  and  $DQ_1G_2^{-1}q = VDQ_1G_2^{-1}q$ . Thus,

$$(DP_1D^-)'(u + DQ_1G_2^{-1}q) = (DP_1D^-)'Uu + (DP_1D^-)'VDQ_1G_2^{-1}q = 0.$$

□

We expect that the solution of the DAE (2.1) should be a function  $x \in C(\mathcal{I}, \mathbb{R}^m)$  that has a continuously differentiable product  $Dx$ , and satisfies the equation at all  $t \in \mathcal{I}$ . We define

$$C_D^1(\mathcal{I}, \mathbb{R}^m) := \{x \in C(\mathcal{I}, \mathbb{R}^m) : Dx \in C^1(\mathcal{I}, \mathbb{R}^n)\}$$

to be the appropriate solution space. We now state the main theorem of März [33] for the finite-dimensional case.

**Theorem 2.1.** *Given a DAE (2.1) with index  $\mu$ ,  $\mu \in \{1, 2\}$ , and  $DP_1D^-$ ,  $DQ_1D^- \in C^1(\mathcal{I}, L(\mathbb{R}^m))$ .*

*Then, the initial value problem for (2.1) with the initial condition*

$$D(t_0)P_1(t_0)(x(t_0) - x^0) = 0, \quad x^0 \in \mathbb{R}^m, \quad (2.18)$$

*and  $t_0 \in \mathcal{I}$ ,  $q \in C_{DQ_1G_2^{-1}}^1(\mathcal{I}, \mathbb{R}^m)$  has a unique solution  $x \in C_D^1(\mathcal{I}, \mathbb{R}^m)$ .*

*Hence, for any*

$$q \in C_{DQ_1G_2^{-1}}^1(\mathcal{I}, \mathbb{R}^m) := \{w \in C(\mathcal{I}, \mathbb{R}^m) : DQ_1G_2^{-1}w \in C^1(\mathcal{I}, \mathbb{R}^n)\},$$

*equation (2.1) is solvable on  $C_D^1(\mathcal{I}, \mathbb{R}^m)$ .*

*Proof.* (From März [33].) Part(i): By standard ODE existence and uniqueness theory, the inherent regular ODE (2.12) with initial condition  $u(t_0) = D(t_0)P_1(t_0)x^0 \in D(t_0)S_1(t_0)$  has a unique solution  $u(t) \in C^1(\mathcal{I}, \mathbb{R}^n)$ . We then use equation (2.10) with  $u(t)$  to construct  $x(t)$ .

In deriving equation (2.10), we had

$$x = D^-DP_1x + D^-DQ_1x + Q_0x.$$

We also know  $u = DP_1x$ . Therefore,

$$Dx = DD^-u + DQ_1x + DQ_0x.$$

However,  $u = Ru = DD^-u$ ,  $DQ_1x = DQ_1G_2^{-1}q$  (see [33]), and  $DQ_0x = D(I - P_0)x = D(I - D^-D)x = Dx - Dx = 0$ . Thus,

$$Dx = u + DQ_1G_2^{-1}q,$$

which is continuously differentiable since  $q \in C^1_{DQ_1G_2^{-1}}(\mathcal{I}, \mathbb{R}^m)$  and therefore we have  $x \in C^1_D(\mathcal{I}, \mathbb{R}^m)$ .

To show uniqueness, we first assume that there is a different  $\hat{x}$  that also solves the DAE (2.1). Letting  $y = x - \hat{x}$  yields the initial condition  $D(t_0)P_1(t_0)y = 0$ . From (2.10) above, we end up with  $y = KD^-DP_1y$ . But the solution  $u = DP_1y$  of the modified inherent regular ODE with  $u(t_0) = 0$  is  $u(t) = 0$ . Thus,  $y = 0$  or  $x = \hat{x}$  and the solution is unique.

That the DAE (2.1) is solvable follows directly from above.  $\square$

*Remark:* For DAEs we need to have *consistent initial conditions*. We can interpret the initial condition given in part (i) of Theorem 2.1 as follows. If we are given an arbitrary condition  $x^0 \in \mathbb{R}^m$ , we can perturb it by using a unique  $x(t_0)$  that satisfies the requirement  $D(t_0)P_1(t_0)(x(t_0) - x^0) = 0$ . We then use this modified consistent initial condition  $x(t_0)$  to guarantee a unique solution to the DAE.

# Chapter 3

## Abstract DAEs: Infinite-dimensional LTI Index-1 Case

### 3.1 Infinite-dimensional Linear Algebra

We first collect some useful infinite-dimensional definitions and statements which will be used in this and the next section. Details and proofs of these statements can be found in standard functional analysis references such as [11], [26], [28] and [38]:

**Proposition 3.1.** *Let  $L(X, Y)$  denote the space of bounded linear operators from the Hilbert space  $X$  into the Hilbert space  $Y$ .  $L(X)$  represents the space  $L(X, X)$ .*

(i) *The finite sum and composition of bounded linear operators is a bounded linear operator. Specifically, for  $T_1, T_2, \dots, T_n \in L(X, Y)$ ,  $K = T_1 + T_2 + \dots + T_n \implies K \in L(X, Y)$  and for  $T_i : X_i \rightarrow X_{i+1}$ ,  $K = T_n T_{n-1} \dots T_2 T_1 \implies K \in L(X_1, X_{n+1})$ .*

(ii) *Given a bounded linear operator  $T \in L(X, Y)$ ,  $\ker T$  is a closed linear subspace, i.e.,  $\ker T = \overline{\ker T}$ .*

(iii) *An operator  $P$  is a projector if it is idempotent, i.e.,  $P^2 = P$ . Furthermore, projectors are bounded i.e.,  $P \in L(X)$ , and both the kernel and image are closed linear subspaces. For the special case of an orthogonal projector, we have  $\text{im } P = \ker(I - P)$  and  $(I - P)^2 = (I - P) \in L(X)$  with  $\ker P \oplus \text{im } P = X$ .*

(iv) *A bounded linear operator  $T \in L(X, Y)$  is one-to-one (i.e., injective) if and only if  $\ker T = \{0\}$ .*

(v) *A bounded linear operator  $T \in L(X, Y)$  that is one-to-one and onto all of  $Y$  (i.e., bijective) has a bounded inverse  $T^{-1} : Y \rightarrow X$ .*

- (vi) A projector  $P$  on  $L(X)$  has a finite dimensional image if and only if it is compact.
- (vii) The sum of a closed linear operator and a bounded linear operator is a closed linear operator, i.e., for  $B \in L(X, Y)$  and closed linear operator  $C : X \rightarrow Y$ ,  $B + C$  defined on  $\mathcal{D}(C)$  is a closed linear operator.
- (viii) For  $B \in L(X, Y)$  and closed linear operator  $C : Y \rightarrow Z$ , the composition  $CB$  is a closed linear operator where  $\mathcal{D}(CB) = \{x \in X : Bx \in \mathcal{D}(C)\}$ . In particular, if  $\text{im}(B) \subset \mathcal{D}(C)$ , then  $CB \in L(X, Z)$ .
- (ix) Similarly, for closed linear operator  $C : X \rightarrow Y$  and  $B \in L(Y, Z)$ , the composition  $BC$  is a closed linear operator where  $\mathcal{D}(BC) = \mathcal{D}(C)$ .
- (x) Closed Graph Theorem: A closed linear operator  $C$  defined on all of a Hilbert space  $X$  which maps into a Hilbert space  $Y$  is bounded.
- (xi) If the inverse of a closed operator  $C$  exists, then  $C^{-1}$  is closed.

## 3.2 Abstract Index and Operator Sequence

We will thoroughly examine the linear constant coefficient case also known as the linear time-invariant case.

We define the linear time-invariant abstract DAE: (where prime denotes differentiation with respect to time)

$$\begin{aligned} \mathcal{E}x'(t) + \mathcal{B}x(t) &= q(t), \quad x \in \mathcal{X}, \quad q \in \mathcal{Y}, \quad \text{or,} \\ \mathcal{A}\mathcal{D}x'(t) + \mathcal{B}x(t) &= q(t), \end{aligned} \tag{3.1}$$

where the operator  $\mathcal{E} : \mathcal{X} \rightarrow \mathcal{Y}$  is split or factored into a *well-matched* pair, i.e.,  $\mathcal{E} = \mathcal{A}\mathcal{D}$ , such that  $\ker \mathcal{A} \oplus \text{im } \mathcal{D} = \mathcal{Z}$ . The linear operators defined by  $\mathcal{A} : \mathcal{Z} \rightarrow \mathcal{Y}$ ,  $\mathcal{B} : \mathcal{X} \rightarrow \mathcal{Y}$ ,  $\mathcal{D} : \mathcal{X} \rightarrow \mathcal{Z}$  act on real Hilbert spaces  $\mathcal{X}$ ,  $\mathcal{Y}$ , and  $\mathcal{Z}$ . Furthermore, we will investigate the bounded case where  $\mathcal{E}$ ,  $\mathcal{A}$  and  $\mathcal{D}$  are bounded and normally solvable (i.e., their images are closed). We initially assume  $\mathcal{B}$  is also bounded and densely defined on  $\mathcal{X}$  (i.e., defined on a dense subset  $\mathcal{D}(\mathcal{B}) \subseteq \mathcal{X}$ ). The case where  $\mathcal{B}$  is unbounded is not as interesting but it will provide some insight into how to handle the unbounded, but closed, case that is investigated later.

*Remark.* For  $\mathcal{B}$  bounded, we just have an ordinary differential equation, but on an abstract Hilbert space. Our goal is to work with partial differential equations, so we will need to address the case where  $\mathcal{B}$  is unbounded but closed. We will do so following discussion of the bounded case.

**Definition 3.1.** We generalize the matrix and subspace sequence in Chapter 2 by defining the following iterated time-invariant *operator and subspace sequence*:

$$\begin{aligned}
G_0 &:= \mathcal{E} = \mathcal{A}\mathcal{D}, & B_0 &= \mathcal{B} \\
\text{For } i &= 0, 1 : \\
G_i &: \mathcal{X} \rightarrow \mathcal{Y} \\
\mathfrak{N}_i &:= \ker G_i = \overline{\ker G_i} := \text{closure of } \ker G_i, \\
Q_i &\in L(\mathcal{X}), & Q_i^2 &= Q_i, & \text{im } Q_i &= \mathfrak{N}_i, \\
P_i &:= I - Q_i, \\
W_i &\in L(\mathcal{Y}), & W_i^2 &= W_i, & \ker W_i &= \text{im } G_i = \overline{\text{im } G_i}, \\
\mathfrak{S}_i &= \ker W_i B_i, \\
G_{i+1} &:= G_i + B_i Q_i, & B_{i+1} &:= B_i P_i,
\end{aligned}$$

where all of the above define operators except for  $\mathfrak{N}_i$  and  $\mathfrak{S}_i$  which are subspaces of  $\mathcal{X}$ .

As before, we define  $\mathcal{D}^- : \mathcal{Z} \rightarrow \mathcal{X}$ , the reflexive generalized inverse of  $\mathcal{D}$ , by  $\mathcal{D}^- \mathcal{D} = P_0$  with  $R^2 = R = \mathcal{D}\mathcal{D}^-$ .

**Definition 3.2.** For the general abstract DAE, we define:

- (a) an *abstract index-1 DAE*, if  $\dim(\text{im } W_0) > 0$  and  $G_1$  is injective and densely solvable (i.e.,  $\text{im } G_1$  is dense in  $\mathcal{Y}$ );
- (b) an *abstract index-2 DAE*, if  $\dim(\text{im } W_i) > 0$ ,  $i \in \{0, 1\}$ , and  $G_2$  is injective and densely solvable.

The above statements for an abstract index provide criteria that must be met for an abstract DAE to be index-1 or index-2. As we will see later, these criteria will also help us to ensure that the abstract DAE has a unique solution.

### 3.3 The Index-1 Semi-explicit Case

Let us first begin with the index-1 abstract DAE case of (3.1). We will investigate the index-2 abstract DAE case in the Chapter 4. To be more practical, we first investigate the semi-explicit operator matrix version of this abstract DAE that might arise from the coupling of a system of partial differential equations (PDEs) with an ordinary DAE or

an operator constraint equation.

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \quad (3.2)$$

$$\text{where } \mathcal{E} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\text{and } \mathcal{B} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

Let  $\mathcal{H}_1$ ,  $\mathcal{H}_2$ , and  $\mathcal{H}_3$  be real Hilbert spaces with  $x_1 \in \mathcal{H}_1$ ,  $x_2 \in \mathcal{H}_2$ ,  $q_1 \in \mathcal{H}_1$ , and  $q_2 \in \mathcal{H}_3$ . Then  $\mathcal{E} : \mathcal{H}_1 \times \mathcal{H}_2 \rightarrow \mathcal{H}_1 \times \mathcal{H}_3$  and  $\mathcal{B} : \mathcal{H}_1 \times \mathcal{H}_2 \rightarrow \mathcal{H}_1 \times \mathcal{H}_3$ . We next split  $\mathcal{E}$  into a *well-matched* pair  $\mathcal{E} = \mathcal{A}\mathcal{D}$  such that  $\ker \mathcal{A} \oplus \text{im } \mathcal{D} = \mathcal{H}_1$ . We see that  $\mathcal{A}$  and  $\mathcal{D}$  as defined below satisfies the well-matched criteria.

$$\mathcal{A} = \begin{bmatrix} I \\ 0 \end{bmatrix}$$

$$\mathcal{D} = \begin{bmatrix} I & 0 \end{bmatrix}.$$

Hence, for this special case we have the space  $\mathcal{X} = \mathcal{H}_1 \times \mathcal{H}_2$ , the space  $\mathcal{Y} = \mathcal{H}_1 \times \mathcal{H}_3$ , and the space  $\mathcal{Z} = \mathcal{H}_1$ .

Using the operator sequence from Definition 3.1 above, we next find the projector  $Q_0$  such that  $\text{im } Q_0 = \ker \mathcal{E}$ . Clearly,  $\ker \mathcal{E} = \{0\} \times \mathcal{H}_2$ . Thus, for

$$Q_0 = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}$$

we see that  $Q_0^2 = Q_0$  and that  $\text{im } Q_0 = \{0\} \times \mathcal{H}_2$ . We note that other choices for  $Q_0$  are possible.

From the definitions  $P_0 = I - Q_0$  and  $\mathcal{D}^- \mathcal{D} = P_0$ , we see that

$$P_0 = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{and}$$

$$\mathcal{D}^- = \begin{bmatrix} I \\ 0 \end{bmatrix}.$$

We assume that this linear time-invariant semi-explicit abstract DAE has index - 1. Thus,  $G_1$  must be one-to-one, i.e.,  $\ker G_1 = \{0\} \in \mathcal{H}_1 \times \mathcal{H}_2$ . By definition,  $G_1 = \mathcal{E} + \mathcal{B}Q_0$ , or

$$G_1 = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix},$$

$$G_1 = \begin{bmatrix} I & B_{12} \\ 0 & B_{22} \end{bmatrix}. \quad (3.3)$$



To find  $\ker G_1$  we calculate

$$\begin{bmatrix} I & B_{12} \\ 0 & B_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or,

$$\begin{aligned} x_1 + B_{12}x_2 &= 0 \\ B_{22}x_2 &= 0. \end{aligned}$$

Therefore,

$$\ker G_1 = \begin{bmatrix} -B_{12}x_2 \\ x_2 \end{bmatrix} : x_2 \in \ker B_{22}.$$

However, we want  $\ker G_1 = \{0\}$ , this means that we must satisfy

$$\begin{bmatrix} -B_{12}x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} : x_2 \in \ker B_{22}.$$

We then come to the conclusion that if the semi-explicit DAE (3.2) has index-1, or  $\ker G_1 = \{0\}$ , then we need  $\ker B_{22} = \{0\}$ . Likewise, if  $B_{22}$  is injective, then  $G_1$  is injective and the DAE has index-1.

We have thus proven the following:

**Lemma 3.1.** *The linear time-invariant semi-explicit abstract DAE (3.2) has abstract index-1 if and only if  $B_{22}$  is one-to-one, i.e.,  $\ker B_{22} = \{0\}$ .*

*Remark.* By looking at the representation for  $G_1$  above in (3.3), we observe that if  $B_{12}$  and  $B_{22}$  are bounded operators then  $G_1$  will also be bounded regardless of whether  $B_{11}$  or  $B_{21}$  are bounded.

Since  $B_{22}$  and  $G_1$  are assumed to be invertible for the index-1 case, we can calculate  $G_1^{-1}$  directly:

$$G_1^{-1} = \begin{bmatrix} I & -B_{12}B_{22}^{-1} \\ 0 & B_{22}^{-1} \end{bmatrix}$$

where  $G_1^{-1} : \mathcal{D}(G_1^{-1}) \subset \mathcal{Y} \rightarrow \mathcal{X}$ .

**Lemma 3.2.** *If  $B_{22}$  is bijective, then  $G_1$  is also bijective.*

*Proof.* Given any element  $y = [y_1 \ y_2]^T \in \mathcal{Y}$ , we want to show there exists an  $x = [x_1 \ x_2]^T \in \mathcal{X}$  such that  $G_1x = y$ . We have

$$\begin{bmatrix} I & B_{12} \\ 0 & B_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \text{ or,}$$

$$\begin{aligned}x_1 + B_{12}x_2 &= y_1 \\ B_{22}x_2 &= y_2.\end{aligned}$$

Since  $B_{22}$  is bijective, we can solve for  $x_2$  uniquely in terms of  $y_2$ , i.e.,  $x_2 = B_{22}^{-1}y_2$ . Then, substituting back into the first equation we can solve  $x_1$  uniquely in terms of  $y_1$  and  $y_2$ , that is,  $x_1 = y_1 - B_{12}B_{22}^{-1}y_2$ . Thus, given any  $y \in \mathcal{Y}$ , we can always find an  $x \in \mathcal{X}$  such that  $G_1x = y$ . Hence,  $G_1$  is surjective. From Lemma 3.1, we also have that  $G_1$  is injective. Therefore,  $G_1$  is bijective. We then have  $G_1^{-1} : \mathcal{Y} \xrightarrow{\text{onto}} \mathcal{X}$ . Furthermore, by Proposition 3.1, part (v),  $G_1^{-1}$  is bounded.  $\square$

### 3.4 A Case Where $\mathcal{B}$ is Bounded

Before we state the theorem for the bounded index-1 case, we will first need the following Lemmas:

**Lemma 3.3.** *For  $\mathcal{E}$  and  $\mathcal{B}$  bounded, an index-1 abstract DAE (3.1) can be decoupled into (i) an AODE:*

$$u'(t) + \mathcal{D}G_1^{-1}\mathcal{B}\mathcal{D}^{-1}u(t) = \mathcal{D}G_1^{-1}q(t), \text{ where } u(t) = \mathcal{D}x(t), \quad (3.4)$$

(ii) and constraint equation:

$$Q_0(I + G_1^{-1}\mathcal{B}P_0)x(t) = Q_0G_1^{-1}q(t). \quad (3.5)$$

(iii) Furthermore, if  $u(t)$  is a solution to the AODE (3.4) above, then we have the following solution representation formula for (3.1):

$$x(t) = (I - Q_0G_1^{-1}\mathcal{B})\mathcal{D}^{-1}u(t) + Q_0G_1^{-1}q(t). \quad (3.6)$$

*Proof.* The development of the inherent abstract ODE, the constraint equation and the solution representation parallels that given in Balla & März [2] for the index-1 finite dimensional case. We will prove the Lemma for the general DAE (3.1) with  $\mathcal{E} = \mathcal{A}\mathcal{D}$  a well-matched pair:

$$\mathcal{A}\mathcal{D}x'(t) + \mathcal{B}x(t) = q(t).$$

We use the properties of the reflexive generalized inverse for  $\mathcal{D} = \mathcal{D}\mathcal{D}^{-1}\mathcal{D}$  and for the orthogonal projectors  $P_0 + Q_0 = I$  to rewrite the above DAE as:

$$\mathcal{A}\mathcal{D}\mathcal{D}^{-1}\mathcal{D}x'(t) + \mathcal{B}(P_0 + Q_0)x(t) = q(t).$$

We see that  $\mathcal{B}Q_0\mathcal{D}^-\mathcal{D}x'(t) = 0$  and  $\mathcal{A}\mathcal{D}Q_0x(t) = 0$  since  $Q_0\mathcal{D}^-\mathcal{D} = Q_0P_0 = 0$  and  $\mathcal{D}Q_0 = \mathcal{D}\mathcal{D}^-\mathcal{D}Q_0 = \mathcal{D}P_0Q_0 = 0$ . We can therefore add the two zero terms to the equation above and then factor to obtain

$$\begin{aligned} \mathcal{A}\mathcal{D}\mathcal{D}^-\mathcal{D}x'(t) + \mathcal{B}Q_0\mathcal{D}^-\mathcal{D}x'(t) + \mathcal{A}\mathcal{D}Q_0x(t) + \mathcal{B}Q_0x(t) + \mathcal{B}P_0x(t) &= q(t) \\ (\mathcal{A}\mathcal{D} + \mathcal{B}Q_0)\{\mathcal{D}^-\mathcal{D}x'(t) + Q_0x(t)\} + \mathcal{B}P_0x(t) &= q(t) \\ G_1\{\mathcal{D}^-\mathcal{D}x'(t) + Q_0x(t)\} + \mathcal{B}P_0x(t) &= q(t). \end{aligned}$$

In the above, we used the identities  $G_1 = \mathcal{A}\mathcal{D} + \mathcal{B}Q_0$  and  $Q_0^2 = Q_0$ . Because this is an index-1 DAE with bounded  $\mathcal{B}$ ,  $G_1$  is bijective. Therefore, multiplying the end result above by  $G_1^{-1}$  is well-defined. This yields

$$\mathcal{D}^-\mathcal{D}x'(t) + Q_0x(t) + G_1^{-1}\mathcal{B}P_0x(t) = G_1^{-1}q(t). \quad (3.7)$$

We can then decouple the above equation (3.7) by multiplying by the orthogonal projectors  $P_0$  and  $Q_0$  to obtain two equations. Multiplying first by  $P_0 = \mathcal{D}^-\mathcal{D}$  gives us

$$\begin{aligned} \mathcal{D}^-\mathcal{D}\mathcal{D}^-\mathcal{D}x'(t) + \mathcal{D}^-\mathcal{D}Q_0x(t) + \mathcal{D}^-\mathcal{D}G_1^{-1}\mathcal{B}P_0x(t) &= \mathcal{D}^-\mathcal{D}G_1^{-1}q(t) \\ \mathcal{D}^-\mathcal{D}x'(t) + \mathcal{D}^-\mathcal{D}Q_0x(t) + \mathcal{D}^-\mathcal{D}G_1^{-1}\mathcal{B}P_0x(t) &= \mathcal{D}^-\mathcal{D}G_1^{-1}q(t). \end{aligned}$$

We now multiply both sides by  $\mathcal{D}$  to put the image of the equation into the space  $\mathcal{Z}$  which yields

$$\begin{aligned} \mathcal{D}\mathcal{D}^-\mathcal{D}x'(t) + \mathcal{D}\mathcal{D}^-\mathcal{D}Q_0x(t) + \mathcal{D}\mathcal{D}^-\mathcal{D}G_1^{-1}\mathcal{B}P_0x(t) &= \mathcal{D}\mathcal{D}^-\mathcal{D}G_1^{-1}q(t) \\ (\mathcal{D}x)'(t) + \mathcal{D}Q_0x(t) + \mathcal{D}G_1^{-1}\mathcal{B}\mathcal{D}^-(\mathcal{D}x)(t) &= \mathcal{D}G_1^{-1}q(t). \end{aligned}$$

But  $\mathcal{D}Q_0 = 0$  and we define  $u(t) := \mathcal{D}x(t) \in \mathcal{Z}$  to obtain the *Abstract Ordinary Differential Equation* (AODE):

$$u'(t) + \mathcal{D}G_1^{-1}\mathcal{B}\mathcal{D}^-u(t) = \mathcal{D}G_1^{-1}q(t). \quad (3.8)$$

This completes the proof of part (i).

We next multiply equation (3.7) by  $Q_0$  to obtain

$$Q_0\mathcal{D}^-\mathcal{D}x'(t) + Q_0^2x(t) + Q_0G_1^{-1}\mathcal{B}P_0x(t) = Q_0G_1^{-1}q(t).$$

As before, we have  $Q_0\mathcal{D}^-\mathcal{D} = 0$  and  $Q_0^2 = Q_0$ . We have now derived the *constraint* equation completing part (ii):

$$\begin{aligned} Q_0x(t) + Q_0G_1^{-1}\mathcal{B}P_0x(t) &= Q_0G_1^{-1}q(t) \\ Q_0(I + G_1^{-1}\mathcal{B}P_0)x(t) &= Q_0G_1^{-1}q(t). \end{aligned} \quad (3.9)$$

We next derive a representation of the solution for the general linear time-invariant, index-1 abstract DAE (3.1). We have

$$\begin{aligned}
x(t) &= (P_0 + Q_0)x(t) = P_0x(t) + Q_0x(t) \\
&= \mathcal{D}^- \mathcal{D}x(t) + Q_0x(t) \\
&= \mathcal{D}^- \mathcal{D}x(t) - Q_0G_1^{-1}\mathcal{B}P_0x(t) + Q_0G_1^{-1}q(t) \\
&= \mathcal{D}^- \mathcal{D}x(t) - Q_0G_1^{-1}\mathcal{B}\mathcal{D}^- \mathcal{D}x(t) + Q_0G_1^{-1}q(t) \\
&= (I - Q_0G_1^{-1}\mathcal{B})\mathcal{D}^- (\mathcal{D}x)(t) + Q_0G_1^{-1}q(t).
\end{aligned}$$

In the above we substituted for  $Q_0x$  from the first part of equation (3.9). Since we assumed existence of a solution  $u(t)$  for the AODE (3.4), we then substitute  $u(t) = \mathcal{D}x(t)$  into the above to obtain the general solution representation:

$$x(t) = (I - Q_0G_1^{-1}\mathcal{B})\mathcal{D}^- u(t) + Q_0G_1^{-1}q(t), \quad (3.10)$$

where  $u(t)$  satisfies the AODE (3.8).

Now we need to verify that (3.10) actually satisfies the abstract DAE (3.1). We do this by verifying it satisfies the equivalent equation (3.7). In the following we use the previously established identity  $\mathcal{D}Q_0 = Q_0\mathcal{D}^- = 0$ . We substitute the solution representation (3.10) into the equation (3.7), which yields

$$\begin{aligned}
&\mathcal{D}^- \mathcal{D}x'(t) + Q_0x(t) + G_1^{-1}\mathcal{B}P_0x(t) \\
&= \mathcal{D}^- \mathcal{D}[(I - Q_0G_1^{-1}\mathcal{B})\mathcal{D}^- u(t) + Q_0G_1^{-1}q(t)]' \\
&\quad + Q_0[(I - Q_0G_1^{-1}\mathcal{B})\mathcal{D}^- u(t) + Q_0G_1^{-1}q(t)] \\
&\quad + G_1^{-1}\mathcal{B}P_0[(I - Q_0G_1^{-1}\mathcal{B})\mathcal{D}^- u(t) + Q_0G_1^{-1}q(t)] \\
&= [\mathcal{D}^- \mathcal{D}\mathcal{D}^- u(t) - \mathcal{D}^- \mathcal{D}Q_0G_1^{-1}\mathcal{B}\mathcal{D}^- u(t) + \mathcal{D}^- \mathcal{D}Q_0G_1^{-1}q(t)]' \\
&\quad + Q_0\mathcal{D}^- u(t) - Q_0G_1^{-1}\mathcal{B}\mathcal{D}^- u(t) + Q_0G_1^{-1}q(t) + G_1^{-1}\mathcal{B}P_0\mathcal{D}^- u(t) \\
&\quad - G_1^{-1}\mathcal{B}P_0Q_0G_1^{-1}\mathcal{B}\mathcal{D}^- u(t) + G_1^{-1}\mathcal{B}P_0Q_0G_1^{-1}q(t) \\
&= [\mathcal{D}^- u(t)]' - Q_0G_1^{-1}\mathcal{B}\mathcal{D}^- u(t) + Q_0G_1^{-1}q(t) + G_1^{-1}\mathcal{B}P_0\mathcal{D}^- u(t) \\
&= \mathcal{D}^- u'(t) - (I - P_0)G_1^{-1}\mathcal{B}\mathcal{D}^- u(t) + Q_0G_1^{-1}q(t) + G_1^{-1}\mathcal{B}P_0\mathcal{D}^- u(t) \\
&= \mathcal{D}^- u'(t) - G_1^{-1}\mathcal{B}\mathcal{D}^- u(t) + \mathcal{D}^- \mathcal{D}G_1^{-1}\mathcal{B}\mathcal{D}^- u(t) + Q_0G_1^{-1}q(t) + G_1^{-1}\mathcal{B}P_0\mathcal{D}^- u(t) \\
&= \mathcal{D}^- [u'(t) + \mathcal{D}G_1^{-1}\mathcal{B}\mathcal{D}^- u(t)] + Q_0G_1^{-1}q(t) \\
&= \mathcal{D}^- [\mathcal{D}G_1^{-1}q(t)] + Q_0G_1^{-1}q(t) \\
&= P_0G_1^{-1}q(t) + Q_0G_1^{-1}q(t) = (P_0 + Q_0)G_1^{-1}q(t) \\
&= G_1^{-1}q(t).
\end{aligned}$$

Thus, the solution representation (3.10) for  $x(t)$  does in fact satisfy (3.7) and hence the abstract index-1 DAE (3.1).  $\square$

*Remark.* We note that in this proof we also showed implicitly that the solution representation for  $x(t)$  is differentiable with respect to time.

*Remark.* The AODE established for this Lemma is for the linear time-invariant, abstract index-1 case only. The time-varying and/or abstract index-2 AODE would be different. See, for example, Chapter 4 for the LTI index-2 abstract Hessenberg case.

*Remark.* In the expression for the solution representation,  $x(t)$ , we can show  $I - Q_0 G_1^{-1} \mathcal{B} = P_{0c}$  where  $P_{0c}$  is the canonical projector onto  $\mathfrak{S}_0$  along  $\mathfrak{N}_0$ . Thus,  $P_{0c}$  is uniquely defined by the given data of the problem. This can be seen from  $P_{0c} = I - Q_{0c}$  where  $Q_{0c}$  can be obtained from the expression  $Q_{0c} = Q_0 G_1^{-1} \mathcal{B}$  for  $Q_0$  any projection onto  $\mathfrak{N}_0$  and  $G_1^{-1}$  derived from that  $Q_0$ . This can be shown in a similar manner as Proposition 2.1.

As remarked earlier, there is more than one choice for the projector  $Q_0$ . This directly leads to a different result for  $P_0$  and ultimately  $\mathcal{D}^-$ . How does a different choice for  $Q_0$  affect the resulting AODE that we derive from the index-1 DAE and the corresponding solution representation? The next Lemma addresses this issue.

**Lemma 3.4.** *The index-1 AODE and solution representation in Lemma 3.3 are unique in that they depend only on the given data for the problem, i.e., the operators  $\mathcal{E}$  and  $\mathcal{B}$ , and not on the choice of projector  $Q_0$ .*

*Proof.* We will extend the proof of Balla & März [2]. First, we claim that the operator compositions  $\mathcal{D}G_1^{-1}$  and  $\mathcal{D}G_1^{-1}\mathcal{B}\mathcal{D}^-$  in the AODE are not affected by our choice of  $P_0$ .

To see this, let  $P_0$  and  $\tilde{P}_0$  be two projectors along  $\mathfrak{N}_0 = \ker \mathcal{E} = \ker \mathcal{A}\mathcal{D}$  that arise from two different choices of  $Q_0$ . We then determine the respective reflexive generalized inverses  $\mathcal{D}^-$  and  $\tilde{\mathcal{D}}^-$ . We next note that the projector  $R : Z \rightarrow Z$  is uniquely defined by the factoring  $\mathcal{E} = \mathcal{A}\mathcal{D}$  since  $R$  is the projection onto  $\text{im } \mathcal{D}$  along  $\ker \mathcal{A}$ . Since  $R$  is unique, we have  $R = \mathcal{D}\mathcal{D}^- = \mathcal{D}\tilde{\mathcal{D}}^-$ . We also have the identity  $\tilde{Q}_0 = Q_0\tilde{Q}_0$  which implies

$$\begin{aligned} \tilde{G}_1 &= \mathcal{A}\mathcal{D} + \mathcal{B}\tilde{Q}_0 \\ &= \mathcal{A}\mathcal{D} + \mathcal{B}Q_0\tilde{Q}_0 \\ &= \mathcal{A}\mathcal{D}\mathcal{D}^- \mathcal{D} + \mathcal{A}\mathcal{D}\tilde{Q}_0 + \mathcal{B}Q_0P_0 + \mathcal{B}Q_0\tilde{Q}_0 \\ &= (\mathcal{A}\mathcal{D} + \mathcal{B}Q_0)(P_0 + \tilde{Q}_0) \\ &= G_1(P_0 + \tilde{Q}_0). \end{aligned}$$

Thus,  $\tilde{G}_1 = G_1(P_0 + \tilde{Q}_0)$ . We use this identity to get an expression for  $\tilde{G}_1^{-1}$ .

$$\begin{aligned}
\tilde{G}_1 &= G_1(P_0 + \tilde{Q}_0) \\
I &= G_1(P_0 + \tilde{Q}_0)\tilde{G}_1^{-1} \\
G_1^{-1} &= (P_0 + \tilde{Q}_0)\tilde{G}_1^{-1} \\
(\tilde{P}_0 + Q_0)G_1^{-1} &= (\tilde{P}_0 + Q_0)(P_0 + \tilde{Q}_0)\tilde{G}_1^{-1} \\
(\tilde{P}_0 + Q_0)G_1^{-1} &= (\tilde{P}_0P_0 + \tilde{P}_0\tilde{Q}_0 + Q_0P_0 + Q_0\tilde{Q}_0)\tilde{G}_1^{-1} \\
(\tilde{P}_0 + Q_0)G_1^{-1} &= (\tilde{P}_0 + \tilde{Q}_0)\tilde{G}_1^{-1} \\
(\tilde{P}_0 + Q_0)G_1^{-1} &= \tilde{G}_1^{-1}.
\end{aligned}$$

Finally, we see

$$\begin{aligned}
\mathcal{D}\tilde{G}_1^{-1} &= \mathcal{D}(\tilde{P}_0 + Q_0)G_1^{-1} \\
\mathcal{D}\tilde{G}_1^{-1} &= \mathcal{D}\tilde{P}_0G_1^{-1} + \mathcal{D}Q_0G_1^{-1} \\
\mathcal{D}\tilde{G}_1^{-1} &= \mathcal{D}\tilde{\mathcal{D}}^-\mathcal{D}G_1^{-1} \\
\mathcal{D}\tilde{G}_1^{-1} &= \mathcal{D}\mathcal{D}^-\mathcal{D}G_1^{-1} \\
\mathcal{D}\tilde{G}_1^{-1} &= \mathcal{D}G_1^{-1}.
\end{aligned}$$

Thus, we have  $\mathcal{D}\tilde{G}_1^{-1} = \mathcal{D}G_1^{-1}$ .

Furthermore,

$$\begin{aligned}
\mathcal{D}\tilde{G}_1^{-1}\mathcal{B}\tilde{\mathcal{D}}^- &= \mathcal{D}\tilde{G}_1^{-1}\mathcal{B}\tilde{\mathcal{D}}^-\mathcal{D}\tilde{\mathcal{D}}^- = \mathcal{D}\tilde{G}_1^{-1}\mathcal{B}\tilde{\mathcal{D}}^-\mathcal{D}\mathcal{D}^- \\
&= \mathcal{D}\tilde{G}_1^{-1}\mathcal{B}\tilde{P}_0\mathcal{D}^- = \mathcal{D}(\tilde{P}_0 + Q_0)G_1^{-1}\mathcal{B}\tilde{P}_0\mathcal{D}^- \\
&= \mathcal{D}\tilde{P}_0G_1^{-1}\mathcal{B}\tilde{P}_0\mathcal{D}^- + \mathcal{D}Q_0G_1^{-1}\mathcal{B}\tilde{P}_0\mathcal{D}^- \\
&= \mathcal{D}\tilde{\mathcal{D}}^-\mathcal{D}G_1^{-1}\mathcal{B}(I - \tilde{Q}_0)\mathcal{D}^- \\
&= \mathcal{D}G_1^{-1}\mathcal{B}\mathcal{D}^- - \mathcal{D}G_1^{-1}\mathcal{B}\tilde{Q}_0\mathcal{D}^- \\
&= \mathcal{D}G_1^{-1}\mathcal{B}\mathcal{D}^-,
\end{aligned}$$

where we used the identities  $G_1^{-1}\mathcal{B}\tilde{Q}_0 = \tilde{Q}_0$  and  $\mathcal{D}\tilde{Q}_0 = 0$ . Thus, we have  $\mathcal{D}\tilde{G}_1^{-1}\mathcal{B}\tilde{\mathcal{D}}^- = \mathcal{D}G_1^{-1}\mathcal{B}\mathcal{D}^-$ . We conclude that the AODE is unique for the given problem data.

Next, we address the solution representation,  $x(t)$ , in (3.10). We have the following result:

$$\tilde{Q}_0\tilde{G}_1^{-1} = \tilde{Q}_0(\tilde{P}_0 + Q_0)G_1^{-1} = Q_0G_1^{-1}.$$

We remarked earlier that  $P_{0c} = I - Q_0G_1^{-1}\mathcal{B}$  is unique based on the problem data. Thus,

$\tilde{P}_{0c} = I - \tilde{Q}_0 \tilde{G}_1^{-1} \mathcal{B} = I - Q_0 G_1^{-1} \mathcal{B} = P_{0c}$ . Hence,

$$\begin{aligned} (I - Q_0 G_1^{-1} \mathcal{B}) \mathcal{D}^- &= P_{0c} \mathcal{D}^- = P_{0c} \tilde{P}_0 \mathcal{D}^- \\ &= P_{0c} \tilde{\mathcal{D}}^- \mathcal{D} \mathcal{D}^- = P_{0c} \tilde{\mathcal{D}}^- \mathcal{D} \tilde{\mathcal{D}}^- = P_{0c} \tilde{\mathcal{D}}^- = \tilde{P}_{0c} \tilde{\mathcal{D}}^- \\ &= (I - \tilde{Q}_0 \tilde{G}_1^{-1} \mathcal{B}) \tilde{\mathcal{D}}^-, \end{aligned}$$

where in the second line we used the identity  $P_{0c} = P_{0c} \tilde{P}_0$ . Thus, this solution representation is in fact unique and only depends on the given data for the problem.  $\square$

The final Lemma will be used later to show existence of a unique solution to the AODE.

**Lemma 3.5.** *Given  $\mathcal{L} \in L(\mathcal{Z})$ ,  $\tilde{q} \in C([0, T]; \mathcal{Z})$  and  $u(0) = u_0 \in \mathcal{Z}$  for any  $T$  satisfying  $0 < T < \infty$ , there exists a unique solution  $u \in C^1([0, T]; \mathcal{Z})$  to the inhomogeneous abstract Cauchy problem*

$$\begin{cases} u'(t) + \mathcal{L}u(t) = \tilde{q}(t) \\ u(0) = u_0. \end{cases} \quad (3.11)$$

*Proof.* By Pazy [34], Theorem 1.1.2, since  $\mathcal{L}$  is a bounded linear operator,  $\mathcal{L}$  is the infinitesimal generator of a uniformly continuous semigroup  $\{T(t)\}_{t>0}$ , and  $T(t) \in L(\mathcal{Z})$  for each  $t > 0$ . Furthermore,  $T(t) = e^{t\mathcal{L}} = \sum_{n=1}^{\infty} \frac{(t\mathcal{L})^n}{n!}$  which converges and is, therefore, well-defined on the Hilbert space  $\mathcal{Z}$ .

There then exists a solution  $u(t) \in \mathcal{Z}$  to the inhomogeneous abstract Cauchy problem (3.11), using a ‘‘variation of parameters’’ formula:

$$u(t) = e^{t\mathcal{L}} u_0 + \int_0^t e^{(t-s)\mathcal{L}} \tilde{q}(s) ds \quad (3.12)$$

for  $\tilde{q} \in C([0, T]; \mathcal{Z})$ . By the above solution (3.12) it is clear that  $u(t)$  is strongly differentiable. Furthermore, it is easy to see that  $u(t)$  as given by (3.12) actually satisfies the inhomogeneous Cauchy problem including the initial condition  $u(0) = u_0$ .

To show uniqueness, we use the standard procedure. Let  $v(t)$  be a second solution to the inhomogeneous abstract Cauchy problem with initial condition  $v(0) = u_0$ . Define  $w(t) = u(t) - v(t)$ . Then,  $w(t)$  satisfies the homogeneous abstract Cauchy problem

$$\begin{cases} w'(t) + \mathcal{L}w(t) = 0 \\ w(0) = 0. \end{cases}$$

This problem has the single solution  $w(t) \equiv 0$ . Hence, we have  $v(t) = u(t)$  and the solution is unique.  $\square$

We then specify some terminology that will be used.

**Definition 3.3.** The initial condition  $x_0$  is called a *consistent* index-1 initial condition if  $x(t_0)$  also satisfies the constraint equation (3.5) for the index-1 abstract DAE.

We can now state the following Theorem.

**Theorem 3.1.** *Given the following: (i)  $\mathcal{B}$  bounded with  $B_{22}$  bijective, (ii) the right-hand side function  $q \in C([0, T]; \mathcal{Y})$ , and (iii) a consistent index-1 initial condition  $x(t_0) \in \mathcal{X}$ ,  $0 < t_0 < T$ ,  $0 < T < \infty$ .*

*Then there exists a unique solution  $x \in C([0, T]; \mathcal{X})$  to the index-1, semi-explicit, abstract DAE (3.2).*

*Proof.* By Lemma 3.2, since  $B_{22}$  is bijective,  $G_1$  is bijective. Therefore, we have the AODE (3.4), constraint equation (3.5) and solution representation (3.6) from Lemma 3.3.

Our plan is as follows. If we can show there exists a unique solution to the AODE (3.4), then by (3.6) there will exist a solution to the general linear time-invariant, abstract index-1 DAE (3.1). We then will need to show uniqueness of that solution to the general DAE.

We now return to the AODE (3.4). First of all, since  $\mathcal{D} : \mathcal{X} \rightarrow \mathcal{Z}$ ,  $\mathcal{D}^- : \mathcal{Z} \rightarrow \mathcal{X}$ ,  $\mathcal{B} : \mathcal{X} \rightarrow \mathcal{Y}$ , and  $G_1^{-1} : \mathcal{Y} \rightarrow \mathcal{X}$ , we see that  $u(t) \in \mathcal{Z}$ . By hypothesis, we are given an initial condition  $x(t_0)$  for the DAE (3.1). As  $u(t) = \mathcal{D}x(t)$ , we get the resulting initial condition for the AODE:  $u(t_0) = \mathcal{D}x(t_0) \in \mathcal{Z}$ .

Since  $\mathcal{E}$  and  $\mathcal{B}$  are bounded, the operators  $\mathcal{D}$ ,  $\mathcal{D}^-$  and  $G_1^{-1}$  are bounded. By Proposition 3.1, part (i), the composition of bounded linear operators is a bounded linear operator. Therefore, we can define the operator  $\mathcal{L} \in L(\mathcal{Z})$  such that  $\mathcal{L} = \mathcal{D}G_1^{-1}\mathcal{B}\mathcal{D}^-$ . We also define  $\tilde{q}(t) \in \mathcal{Z}$  by  $\tilde{q}(t) = \mathcal{D}G_1^{-1}q(t)$ . We rewrite the AODE (3.4) as

$$u'(t) + \mathcal{L}u(t) = \tilde{q}(t). \quad (3.13)$$

Note that  $q \in C([t_0, T]; \mathcal{Y})$  implies  $\tilde{q} \in C([t_0, T]; \mathcal{Z})$  and we have

$$\begin{aligned} \int_{t_0}^T \tilde{q}(s) ds &= \int_{t_0}^T \mathcal{D}G_1^{-1}q(s) ds \\ &\leq \|\mathcal{D}\| \|G_1^{-1}\| \int_{t_0}^T q(s) ds < \infty. \end{aligned}$$

We now invoke Lemma 3.5 which states that the AODE (3.4) has a unique solution  $u(t)$ . By Lemma 3.4, this AODE is uniquely defined by the given data and is not dependent on a particular choice of  $Q_0$ .



Since we have existence of a unique solution  $u(t)$  which satisfies the AODE (3.4), we can write a solution representation for  $x(t)$  using (3.6). We therefore have existence of a solution to the DAE. Again, by Lemma 3.4, this solution representation is uniquely defined by the given data and is not dependent on a particular choice of  $Q_0$ .

We now verify uniqueness of the solution to the DAE (3.2) with initial condition  $\mathcal{D}x(t_0) = f$  where  $f \in \mathcal{Z}$ . If  $\tilde{x}(t)$  is a second different solution in addition to  $x(t)$  for the DAE, then  $y(t) = \tilde{x}(t) - x(t)$  is a solution to the modified homogeneous DAE  $\mathcal{A}\mathcal{D}y'(t) + \mathcal{B}y(t) = 0$  with initial condition  $\mathcal{D}y(t_0) = 0$ . But by the AODE (3.13) with  $\tilde{q}(t) = 0$ ,  $u(t) = \mathcal{D}y(t)$  and  $u(t_0) = \mathcal{D}y(t_0) = 0$ , we get the solution  $u(t) = \mathcal{D}y(t) = 0$ . Thus,  $P_0y(t) = \mathcal{D}^-\mathcal{D}y(t) = 0$  and by the constraint equation (3.5) we see  $Q_0y(t) = 0$  also. Therefore, we have  $\tilde{x}(t) = x(t) + y(t) = x(t) + P_0y(t) + Q_0y(t) = x(t)$  which is a contradiction of our original assumption. Therefore, the solution  $x(t)$  is unique.

By the solution representation (3.6), for  $u(t) \in C^1([0, T]; \mathcal{Z})$  and  $q(t) \in C([0, T]; \mathcal{Y})$ , we have that  $x(t) \in C([0, T]; \mathcal{X})$ .  $\square$

*Remark.* Instead of all the work that was required to show existence and uniqueness for the index-1 semi-explicit DAE (3.2) using this projection framework, we could have done so using a *direct method*. Specifically, the semi-explicit DAE can be written as two equations:

$$\begin{aligned} x_1'(t) + B_{11}x_1(t) + B_{12}x_2(t) &= q_1(t), \\ B_{21}x_1(t) + B_{22}x_2(t) &= q_2(t). \end{aligned}$$

Since  $B_{22}$  is bijective, we can solve for  $x_2(t)$  in the second equation to obtain  $x_2(t) = B_{22}^{-1}q_2(t) - B_{22}^{-1}B_{21}x_1(t)$ . We substitute for  $x_2(t)$  into the first equation to find  $x_1'(t) + (B_{11} - B_{12}B_{22}^{-1}B_{21})x_1(t) = q_1(t) - B_{12}B_{22}^{-1}q_2(t)$ . It turns out this abstract ODE is the same as the AODE in (3.4). Thus, if there is a unique solution to the AODE we then have a unique solution for  $x_1(t)$ . We then substitute this into our expression for  $x_2(t) = B_{22}^{-1}q_2(t) - B_{22}^{-1}B_{21}x_1(t)$  to obtain a unique solution for  $x_2(t)$ .

However, the direct method can only be used for index-1 semi-explicit DAEs of the form (3.2), while the projection method described above can be used for a broader range of problems that fit into the more general form of (3.1). Therefore, it was useful to show this method on a known example.

### 3.5 A Case Where $\mathcal{B}$ is Unbounded

In Theorem 3.1 which we just proved above we assumed  $\mathcal{B}$  was bounded. We now generalize to a specific case of an unbounded but closed  $\mathcal{B}$ .

**Theorem 3.2.** *Given the following: (i)  $\mathcal{B}$  contains bounded operators  $B_{12}$ ,  $B_{21}$ , and  $B_{22}$  with  $B_{22}$  bijective, and (ii)  $\mathcal{B}_{11}$  is the infinitesimal generator of a strongly continuous semigroup on  $\mathcal{H}_1$ , (iii) the right-hand side function  $q \in C^1([0, T]; \mathcal{Y})$ , and (iv) a consistent index-1 initial condition  $x(t_0) \in \mathcal{X}$ ,  $0 < t_0 < T$ ,  $0 < T < \infty$ .*

*Then there exists a unique solution  $x \in C([0, T]; \mathcal{X})$  to the index-1, semi-explicit, abstract DAE (3.2).*

*Proof.* Based on our definition of  $\mathcal{B}$  above, since the domain of  $\mathcal{B}_{11}$ ,  $\mathcal{D}(\mathcal{B}_{11})$ , is dense in  $\mathcal{H}_1$ , we see that the domain of  $\mathcal{B}$ ,  $\mathcal{D}(\mathcal{B})$ , is also dense in  $X = \mathcal{H}_1 \times \mathcal{H}_2$ . By definition,  $G_1 = \mathcal{E} + \mathcal{B}Q_0$ . However, for the semi-explicit case, we found  $Q_0$  to be:

$$Q_0 = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}.$$

Since  $\text{im } Q_0 = \{0\} \times \mathcal{H}_2 \subset \mathcal{D}(\mathcal{B})$ , the definition for  $G_1$  makes sense in this case and  $\mathcal{D}(G_1) = \mathcal{X}$ . Since  $\mathcal{B}$  is a closed operator, and by Proposition 3.1, parts (vii), (viii) and (x), we have the composition of closed operator  $\mathcal{B}$  with bounded operator  $Q_0$  is a closed operator. Then, the sum of a bounded operator  $\mathcal{E}$  with closed operator  $\mathcal{B}Q_0$  is a closed operator. Finally, by the Closed Graph Theorem, since  $G_1$  is a closed operator defined on the whole Hilbert space  $\mathcal{X}$ ,  $G_1$  is a bounded operator. However, this can also be seen by equation (3.3) where we calculated  $G_1$  for the semi-explicit case to be:

$$G_1 = \begin{bmatrix} I & B_{12} \\ 0 & B_{22} \end{bmatrix}.$$

Since we assume both  $B_{12}$  and  $B_{22}$  to be bounded, clearly  $G_1$  is also bounded. Furthermore, since  $B_{22}$  is bijective,  $G_1$  is bijective. Hence,  $G_1^{-1}$  exists, is bounded and was found to be:

$$G_1^{-1} = \begin{bmatrix} I & -B_{12}B_{22}^{-1} \\ 0 & B_{22}^{-1} \end{bmatrix}.$$

Since  $G_1^{-1}$  is bounded, it can be extended to be defined on the entire Hilbert space  $\mathcal{Y}$ .

As noted above and by Proposition 3.1, part (viii),  $\mathcal{B}Q_0$  is actually a bounded operator since  $\text{im } (Q_0) \subset \mathcal{D}(\mathcal{B})$  and  $\mathcal{D}(\mathcal{B}Q_0) = \mathcal{X}$ . Clearly, for  $x \in \mathcal{D}(\mathcal{B})$ ,  $Q_0x \in \mathcal{D}(\mathcal{B})$ . For  $P_0 = I - Q_0$ , we have:

$$P_0 = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}.$$

Likewise for  $x \in \mathcal{D}(\mathcal{B})$ , we have  $P_0x \in \mathcal{D}(\mathcal{B})$ . Thus, for  $x \in \mathcal{D}(\mathcal{B})$ , the derivation of equation (3.7) in the bounded case is also valid for this version of  $\mathcal{B}$ . We therefore have

$$\mathcal{D}^- \mathcal{D}x'(t) + Q_0x(t) + G_1^{-1} \mathcal{B}P_0x(t) = G_1^{-1}q(t). \quad (3.14)$$

As for the bounded case, we decouple this equation by multiplying by the orthogonal projectors  $P_0$  and  $Q_0$ . We first multiply by  $P_0$ , then by  $\mathcal{D}$ , and finally substitute  $u(t) = \mathcal{D}x(t)$  as before to obtain the *Abstract Ordinary Differential Equation*, or AODE:

$$u'(t) + \mathcal{D}G_1^{-1}\mathcal{B}\mathcal{D}^-u(t) = \mathcal{D}G_1^{-1}q(t). \quad (3.15)$$

As noted the middle term  $\mathcal{D}G_1^{-1}\mathcal{B}\mathcal{D}^-u(t)$  makes sense as  $\mathcal{D}^-u(t) = P_0x(t)$  where we assume  $x \in \mathcal{D}(\mathcal{B})$  and hence  $P_0x \in \mathcal{D}(\mathcal{B})$ . Also  $u(t) = \mathcal{D}x(t) = x_1(t)$ , where  $x_1 \in \mathcal{D}(\mathcal{B}_{11}) \subset \mathcal{H}_1$ . If we multiply out the various operator matrices for the specific semi-explicit case for  $\mathcal{D}$ ,  $G_1^{-1}$ ,  $\mathcal{B}$ , and  $\mathcal{D}^-$ , we have

$$u'(t) + (\mathcal{B}_{11} - B_{12}B_{22}^{-1}B_{21})u(t) = q_1(t) - B_{12}B_{22}^{-1}q_2(t). \quad (3.16)$$

We note that the composition  $B_{12}B_{22}^{-1}B_{21}$  is a bounded operator since we assumed  $B_{12}$ ,  $B_{21}$ , and  $B_{22}$  are bounded and by Proposition 3.1, part (i), the composition of bounded operators is a bounded operator. Then,  $(\mathcal{B}_{11} - B_{12}B_{22}^{-1}B_{21})$  becomes a bounded perturbation to an infinitesimal generator of a strongly continuous semigroup. We can therefore apply standard perturbation theory to give us our desired result, see Pazy Section 3.1 [34]:

**Lemma 3.6.** *Let  $\mathcal{A}$  be the infinitesimal generator of a strongly continuous semigroup on a Hilbert space  $\mathcal{X}$ . If  $\mathcal{B}$  is a bounded linear operator,  $\mathcal{B} \in L(\mathcal{X}, \mathcal{Y})$ , then  $\mathcal{A} + \mathcal{B}$  is the infinitesimal generator of a strongly continuous semigroup on  $\mathcal{X}$ .*

Thus, by Lemma 3.6, the sum  $(\mathcal{B}_{11} - B_{12}B_{22}^{-1}B_{21})$  is also the infinitesimal generator of a strongly continuous semigroup in  $\mathcal{H}_1$ . Furthermore,  $\mathcal{D}(\mathcal{B}_{11} - B_{12}B_{22}^{-1}B_{21}) = \mathcal{D}(\mathcal{B}_{11})$  is dense in  $\mathcal{H}_1$ .

We are now ready to state a uniqueness and existence result for the AODE (3.16) that follows from Pazy Section 4.2 [34]:

**Lemma 3.7.** *Let  $\mathcal{L}$  be the infinitesimal generator of a strongly continuous semigroup  $T(t)$  on a Hilbert space  $\mathcal{X}$ . If  $f$  is Lipschitz continuous on the interval  $[0, T]$  then the initial value problem  $u'(t) + \mathcal{L}u(t) = f(t)$ ,  $t > 0$ ,  $u(0) = x$  has a unique solution  $u$  on  $[0, T]$  for every  $x \in \mathcal{D}(\mathcal{L})$ . Furthermore, this solution is given by*

$$u(t) = T(t)x + \int_0^t T(t-s)f(s)ds. \quad (3.17)$$

We let  $\mathcal{L} = (\mathcal{B}_{11} - B_{12}B_{22}^{-1}B_{21})$  and  $f(t) = \mathcal{D}G_1^{-1}q(t) = q_1(t) - B_{12}B_{22}^{-1}q_2(t)$ , then by Lemma 3.7 there is a unique solution  $u(t)$  to the AODE (3.16).

The rest of the proof follows from Theorem 3.1 with only minor modifications from the bounded case.  $\square$

*Remark.* Using other perturbation results for semigroups, we can easily modify Theorem 3.2. For example, we use the following Lemma from Pazy Section 3.3 [34] to obtain a Corollary.

**Lemma 3.8.** *Let  $\mathcal{A}$  be the infinitesimal generator of a strongly continuous semigroup of contractions in a Hilbert space  $\mathcal{X}$ . Let  $\mathcal{B}$  be dissipative such that  $\mathcal{D}(\mathcal{B}) \supset \mathcal{D}(\mathcal{A})$  and*

$$\|\mathcal{B}x\| \leq \|\mathcal{A}x\| + b\|x\| \quad \text{for } x \in \mathcal{D}(\mathcal{A})$$

where  $b \geq 0$ . Then the closure of  $\mathcal{A} + \mathcal{B}$  is the infinitesimal generator of a strongly continuous semigroup of contractions in  $\mathcal{X}$ .

Using this Lemma we can now modify the theorem.

**Corollary 3.1.** *For the conditions in Theorem 3.2 where the matrix operator  $\mathcal{B}$  consists of the elements:  $\mathcal{B}_{11}$  the infinitesimal generator of a strongly continuous semigroup of contractions;  $B_{12}$  and  $B_{22}$  linear bounded operators with  $B_{22}$  bijective; and  $\mathcal{B}_{21}$  a linear closed unbounded operator such that  $B_{12}B_{22}^{-1}\mathcal{B}_{21}$  satisfies the conditions for the operator  $\mathcal{B}$  above in Lemma 3.8. Then, there exists a unique solution to the (index-1) semi-explicit abstract DAE (3.2).*

We note that many modifications to the above Corollary can be made. The most obvious one is for the case when  $\mathcal{L} = (\mathcal{B}_{11} - \mathcal{B}_{12}\mathcal{B}_{22}^{-1}\mathcal{B}_{21})$  is the infinitesimal generator of a strongly continuous semigroup.

**Corollary 3.2.** *Given the following: (i) The combination  $(\mathcal{B}_{11} - \mathcal{B}_{12}\mathcal{B}_{22}^{-1}\mathcal{B}_{21})$  is an infinitesimal generator of a strongly continuous semigroup, (ii) the right-hand side function  $q \in C^1([0, T]; \mathcal{Y})$ , and (iii) a consistent index-1 initial condition  $x(t_0) \in \mathcal{X}$ ,  $0 < t_0 < T$ ,  $0 < T < \infty$ .*

*Then there exists a unique solution  $x \in C([0, T]; \mathcal{X})$  to the index-1, semi-explicit, abstract DAE (3.2).*

# Chapter 4

## Abstract DAEs: Infinite-dimensional LTI Index-2 Case

### 4.1 Background for Abstract Hessenberg DAE

As for the index-1 case, we will also investigate the semi-explicit operator matrix version (3.2) for the abstract index-2 case. From Lemma 3.1 we saw that the abstract semi-explicit DAE had index-1 when the operator  $B_{22}$  was bijective. Since we want to investigate the index-2 case, we will first assume the opposite extreme that  $B_{22} = 0$ , i.e.,  $B_{22}$  is the zero operator. This leads us to the abstract Hessenberg DAE (so named for the similarity to the Hessenberg size-2 finite dimensional DAE):

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}, \quad (4.1)$$

with the auxiliary condition that the composition  $B_{21}B_{12}$  is bijective. For the abstract Hessenberg DAE, we will define the spaces  $\mathcal{X} = \mathcal{Y} = \mathcal{H}_1 \times \mathcal{H}_2$  and  $\mathcal{Z} = \mathcal{H}_1$ , where  $\mathcal{H}_1$  and  $\mathcal{H}_2$  are real Hilbert spaces. As we proceed it will become clear why we included an auxiliary condition on the operators  $B_{12}$  and  $B_{21}$ .

We then factor  $\mathcal{E}$  into a well-matched pair  $\mathcal{E} = \mathcal{A}\mathcal{D}$  such that  $\mathcal{A}$ ,  $\mathcal{D}$ ,  $\mathcal{D}^-$ ,  $P_0$ , and  $Q_0$  are all the same as for the index-1 case as these operators only depend on  $\mathcal{E}$ , which is fixed, and not on  $\mathcal{B}$ . Specifically, we have

$$\mathcal{A} = \begin{bmatrix} I \\ 0 \end{bmatrix} \quad \mathcal{D} = [I \ 0]$$
$$Q_0 = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \quad P_0 = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{and } \mathcal{D}^- = \begin{bmatrix} I \\ 0 \end{bmatrix}.$$

For now we will assume  $\mathcal{B}$  is bounded. We then determine  $G_1$  for (4.1). By definition,  $G_1 = \mathcal{E} + \mathcal{B}Q_0$ , or

$$\begin{aligned} G_1 &= \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}, \\ G_1 &= \begin{bmatrix} I & B_{12} \\ 0 & 0 \end{bmatrix}. \end{aligned} \quad (4.2)$$

To find  $\ker G_1$  we note

$$\begin{bmatrix} I & B_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

that is,

$$x_1 + B_{12}x_2 = 0.$$

Therefore,

$$\ker G_1 = \begin{bmatrix} -B_{12}x_2 \\ x_2 \end{bmatrix} : x_2 \in \mathcal{H}_2$$

Thus, for the abstract Hessenberg DAE,  $\ker G_1$  is nontrivial and the abstract DAE is not index-1. We then proceed to calculate  $Q_1$ ,  $P_1$ , and  $G_2$ .

To find  $Q_1$  we use the properties that  $Q_1^2 = Q_1$  and  $\text{im } Q_1 = \ker G_1$ . For  $w = [w_1, w_2]^T \in \mathcal{H}_1 \times \mathcal{H}_2$ , we have for  $\text{im } Q_1$

$$Q_1 w = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} Q_{11}w_1 + Q_{12}w_2 \\ Q_{21}w_1 + Q_{22}w_2 \end{bmatrix}.$$

However  $\text{im } Q_1 = \ker G_1$ , so we have the condition

$$Q_{11}w_1 + Q_{12}w_2 = -B_{12}(Q_{21}w_1 + Q_{22}w_2).$$

This statement is true for all  $w_1$  and  $w_2$ . If we let  $w_1 = 0$ , we have the result  $Q_{12}w_2 = -B_{12}Q_{22}w_2$  or  $Q_{12} = -B_{12}Q_{22}$ . Similarly for  $w_2 = 0$ , we get  $Q_{11} = -B_{12}Q_{21}$ . We therefore write

$$Q_1 = \begin{bmatrix} -B_{12}Q_{21} & -B_{12}Q_{22} \\ Q_{21} & Q_{22} \end{bmatrix}.$$

We next use the fact  $Q_1$  is a projection so  $Q_1^2 = Q_1$ .

$$\begin{aligned} Q_1^2 &= \begin{bmatrix} -B_{12}Q_{21} & -B_{12}Q_{22} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} -B_{12}Q_{21} & -B_{12}Q_{22} \\ Q_{21} & Q_{22} \end{bmatrix} \\ &= \begin{bmatrix} B_{12}Q_{21}B_{12}Q_{21} - B_{12}Q_{22}Q_{21} & B_{12}Q_{21}B_{12}Q_{22} - B_{12}Q_{22}^2 \\ -Q_{21}B_{12}Q_{21} + Q_{22}Q_{21} & -Q_{21}B_{12}Q_{22} + Q_{22}^2 \end{bmatrix} \\ &= \begin{bmatrix} -B_{12}Q_{21} & -B_{12}Q_{22} \\ Q_{21} & Q_{22} \end{bmatrix} = Q_1. \end{aligned}$$

This leads to four equations which can be written as follows:

$$\begin{aligned} B_{12}(Q_{21}B_{12} - Q_{22})Q_{21} &= -B_{12}Q_{21}, \\ B_{12}(Q_{21}B_{12} - Q_{22})Q_{22} &= -B_{12}Q_{22}, \\ (-Q_{21}B_{12} + Q_{22})Q_{21} &= Q_{21}, \\ (-Q_{21}B_{12} + Q_{22})Q_{22} &= Q_{22}. \end{aligned}$$

These four equations all lead to the identity  $Q_{22} - Q_{21}B_{12} = I$  or  $Q_{22} = I + Q_{21}B_{12}$ . Then,

$$Q_1 = \begin{bmatrix} -B_{12}Q_{21} & -B_{12}(I + Q_{21}B_{12}) \\ Q_{21} & I + Q_{21}B_{12} \end{bmatrix}.$$

We note that  $Q_1$  is not unique as we can choose any  $Q_{21} \in L(\mathcal{H}_1, \mathcal{H}_2)$ . For convenience, we choose to let  $Q_{21} = 0$ . Then, we have

$$Q_1 = \begin{bmatrix} 0 & -B_{12} \\ 0 & I \end{bmatrix}. \quad (4.3)$$

As can easily be seen, this  $Q_1$  satisfies in  $Q_1 = \ker G_1$  and  $Q_1^2 = Q_1$ . This also gives us

$$\begin{aligned} P_1 &= I - Q_1 = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} 0 & -B_{12} \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} I & B_{12} \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

We can now calculate  $G_2 = G_1 + \mathcal{B}P_0Q_1$ .

$$\begin{aligned} G_2 &= \begin{bmatrix} I & B_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -B_{12} \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} I & B_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} B_{11} & 0 \\ B_{21} & 0 \end{bmatrix} \begin{bmatrix} 0 & -B_{12} \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} I & B_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -B_{11}B_{12} \\ 0 & -B_{21}B_{12} \end{bmatrix} \\ &= \begin{bmatrix} I & (I - B_{11})B_{12} \\ 0 & -B_{21}B_{12} \end{bmatrix}. \end{aligned} \quad (4.4)$$

We now show that given the auxiliary condition of  $B_{21}B_{12}$  bijective, that  $G_2$  is one-to-one and onto. We first determine  $\ker G_2$ , by considering

$$\begin{bmatrix} I & (I - B_{11})B_{12} \\ 0 & -B_{21}B_{12} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This leads to

$$\begin{aligned}x_1 + (I - B_{11})B_{12}x_2 &= 0 \\ -B_{21}B_{12}x_2 &= 0.\end{aligned}$$

Since  $(B_{21}B_{12})^{-1}$  exists and is defined on all of  $\mathcal{H}_2$ , we get  $x_2 = 0$  and  $x_1 = 0$ . Hence,  $\ker G_2 = \{0\}$  and  $G_2$  is one-to-one. To show  $G_2$  is onto we show that for any  $y = [y_1 \ y_2]^T \in \mathcal{H}_1 \times \mathcal{H}_2$  one can find an  $x = [x_1 \ x_2]^T \in \mathcal{H}_1 \times \mathcal{H}_2$  such that  $G_2x = y$ , i.e.,

$$\begin{bmatrix} I & (I - B_{11})B_{12} \\ 0 & -B_{21}B_{12} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

This implies

$$\begin{aligned}x_1 + (I - B_{11})B_{12}x_2 &= y_1 \\ -B_{21}B_{12}x_2 &= y_2.\end{aligned}$$

Since  $B_{21}B_{12}$  is bijective, we get

$$\begin{aligned}x_1 &= y_1 + (I - B_{11})B_{12}(B_{21}B_{12})^{-1}y_2 \\ x_2 &= -(B_{21}B_{12})^{-1}y_2.\end{aligned}$$

Thus,  $G_2$  is one-to-one and onto.

*Remark.* Since  $G_2$  is one-to-one and onto, the abstract Hessenberg DAE (4.1) (with auxiliary condition of  $B_{21}B_{12}$  bijective) has abstract index-2.

Here we used the  $Q_1$  above to generate a  $G_2$ . However, this  $Q_1$  is not unique in that it can be any projection onto  $\mathfrak{N}_1$ . The direction along which  $Q_1$  projects is not specified, only the space onto which it projects is specified. Thus, there can be many different  $Q_1$ 's that satisfy the conditions of Definition 3.1. This is indicated by the fact that we could choose anything for  $Q_{21}$  in our derivation of  $Q_1$  above. We now use the expression developed in Proposition 2.1 and apply it to the infinite dimensional case. Given any  $Q_1$  which projects onto  $\mathfrak{N}_1$  along with the  $G_2$  derived from that  $Q_1$ , we can find the unique canonical projection  $\tilde{Q}_1$  onto  $\mathfrak{N}_1$  along the space  $\mathfrak{S}_1$  by  $\tilde{Q}_1 = Q_1G_2^{-1}\mathcal{B}P_0$ . Since the  $G_2$  determined above is one-to-one and onto (bijective), a simple calculation shows its inverse to be

$$G_2^{-1} = \begin{bmatrix} I & (I - B_{11})B_{12}(B_{21}B_{12})^{-1} \\ 0 & -(B_{21}B_{12})^{-1} \end{bmatrix}.$$

We then calculate

$$\begin{aligned}\tilde{Q}_1 &= Q_1G_2^{-1}\mathcal{B}P_0 \\ &= \begin{bmatrix} 0 & -B_{12} \\ 0 & I \end{bmatrix} \begin{bmatrix} I & (I - B_{11})B_{12}(B_{21}B_{12})^{-1} \\ 0 & -(B_{21}B_{12})^{-1} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \\ \tilde{Q}_1 &= \begin{bmatrix} B_{12}(B_{21}B_{12})^{-1}B_{21} & 0 \\ -(B_{21}B_{12})^{-1}B_{21} & 0 \end{bmatrix}.\end{aligned}$$



For ease of notation we define  $F = (B_{21}B_{12})^{-1}B_{21}$  and  $H = B_{12}F = B_{12}(B_{21}B_{12})^{-1}B_{21}$ . The operator  $H$  satisfies  $H^2 = H$  and is the projection onto  $\text{im } B_{12}$  along  $\text{ker } B_{21}$ . We also note that  $F$  is a reflexive generalized inverse of  $B_{12}$  since  $FB_{12}F = F$  and  $B_{12}FB_{12} = B_{12}$ .

We can then write

$$Q_1 = \begin{bmatrix} H & 0 \\ -F & 0 \end{bmatrix}. \quad (4.5)$$

We next calculate a new  $G_2$  based on the canonical  $Q_1$ ,

$$\begin{aligned} G_2 &= G_1 + \mathcal{B}P_0Q_1 \\ &= \begin{bmatrix} I & B_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} H & 0 \\ -F & 0 \end{bmatrix} \\ &= \begin{bmatrix} I & B_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} B_{11} & 0 \\ B_{21} & 0 \end{bmatrix} \begin{bmatrix} H & 0 \\ -F & 0 \end{bmatrix} \\ &= \begin{bmatrix} I & B_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} B_{11}H & 0 \\ B_{21}H & 0 \end{bmatrix} \\ &= \begin{bmatrix} I + B_{11}H & B_{12} \\ B_{21}H & 0 \end{bmatrix} \\ &= \begin{bmatrix} I + B_{11}H & B_{12} \\ B_{21} & 0 \end{bmatrix}, \end{aligned} \quad (4.6)$$

where in the last line we used  $B_{21}H = B_{21}B_{12}(B_{21}B_{12})^{-1}B_{21} = B_{21}$ .

As expected this  $G_2$  is also one-to-one and onto. Instead of showing that  $\text{ker } G_2 = \{0\}$  and that given an arbitrary  $y$  we can find an  $x$  such that  $G_2x = y$ , we will go about it in a backwards manner. We will instead exhibit an operator matrix  $G_2^{-1}$  that satisfies the conditions of an inverse and is defined on the entire space. The calculation to derive  $G_2^{-1}$  is not easy. We will use three formulae to derive the inverse:  $G_2^{-1}G_2x = x$ ,  $G_2G_2^{-1}y = y$ , and  $\tilde{Q}_1 = \tilde{Q}_1G_2^{-1}\mathcal{B}P_0$ . The first two formulae are merely the conditions required of an inverse and the latter formula was used to find the canonical projection  $Q_1$  onto  $\mathfrak{N}_1$  along  $\mathfrak{S}_1$ . Clearly, this last formula should be an identity when the canonical projection  $\tilde{Q}_1$  is used along with the  $G_2^{-1}$  derived from that  $\tilde{Q}_1$ .

The details of the calculation are left to the reader. However, it is easy to check that the below operator satisfies the three equations above. The final result is

$$G_2^{-1} = \begin{bmatrix} (I - H) & (I - B_{11} + HB_{11})B_{12}(B_{21}B_{12})^{-1} \\ F & -F(I + HB_{11})B_{12}(B_{21}B_{12})^{-1} \end{bmatrix}.$$

If we let  $T = B_{12}(B_{21}B_{12})^{-1}$  we can write this as

$$G_2^{-1} = \begin{bmatrix} (I - H) & (I - B_{11} + HB_{11})T \\ F & -F(I + HB_{11})T \end{bmatrix}. \quad (4.7)$$

Similarly, we note that  $T$  is a reflexive generalized inverse of  $B_{21}$  since  $TB_{21}T = T$  and  $B_{21}TB_{21} = B_{21}$ . Furthermore, as we might expect  $FT = (B_{21}B_{12})^{-1}$ .

## 4.2 A Case Where $\mathcal{B}$ is Bounded

We next observe that if  $\mathcal{B}$  is bounded, then not only is  $G_2$  bounded but  $G_2^{-1}$  is also bounded. Since  $(B_{21}B_{12})^{-1}$  is one-to-one and onto,  $G_2^{-1}$  is defined on all of  $\mathcal{Y}$ . Hence, this canonical  $G_2$  is also one-to-one and onto. Thus, for  $\mathcal{B}$  bounded and  $(B_{21}B_{12})^{-1}$  bijective, the abstract Hessenberg DAE (4.1) has abstract index-2.

We next write the Hessenberg DAE (4.1) as two equations

$$x_1'(t) + B_{11}x_1(t) + B_{12}x_2(t) = q_1(t) \quad (4.8)$$

$$B_{21}x_1(t) = q_2(t). \quad (4.9)$$

We will also need to define a consistent initial condition for the abstract Hessenberg DAE case.

**Definition 4.1.** The initial condition  $x(t_0)$  is called a *consistent* Hessenberg initial condition if  $x(t_0)$  also satisfies the constraint equation (4.9) for the abstract Hessenberg DAE (4.1).

We can now state the next result.

**Theorem 4.1.** *Given: (i)  $\mathcal{B}$  bounded with the composition  $B_{21}B_{12}$  bijective, (ii) the right-hand side function  $q(t) = [q_1 \ q_2]^T$  satisfies  $q_1 \in C([0, T]; \mathcal{H}_1)$  and  $q_2 \in C^1([0, T]; \mathcal{H}_2)$ , and (iii) a consistent Hessenberg initial condition  $x_1(0)$ .*

*Then there exists a unique solution  $x \in C([0, T]; \mathcal{X})$  to the index-2 semi-explicit abstract Hessenberg DAE (4.1).*

*Proof.* Since  $\mathcal{B}$  is bounded and  $(B_{21}B_{12})^{-1}$  is bijective, we know from our calculations above that  $G_2^{-1}$  exists on the entire space and is bounded. Thus, the abstract DAE (4.1) has abstract index-2. As in the index-1 case, we now work to decouple the abstract Hessenberg DAE (4.1) into an inherent abstract ODE (AODE), its constraint equation, and a solution representation. As in the finite-dimensional case (see [33]), we have  $\mathcal{D}P_1\mathcal{D}^- = I - H$  and  $\mathcal{D}Q_1\mathcal{D}^- = H$  are projectors that yield the respective decomposition  $\mathcal{D}\mathcal{S}_1 \oplus \mathcal{D}\mathcal{N}_1 = \text{im } \mathcal{D} = \mathcal{Z} = \mathcal{H}_1$ . We multiply the second equation (4.9) above by  $T$ , the reflexive generalized inverse of  $B_{21}$ , where  $T = B_{12}(B_{21}B_{12})^{-1}$  to obtain

$$Hx_1 = Tq_2. \quad (4.10)$$

We then decouple the first equation (4.8) into two equations by multiplying first by  $F$ , the reflexive generalized inverse of  $B_{12}$ , where  $F = (B_{21}B_{12})^{-1}B_{21}$  and second by  $\mathcal{D}P_1\mathcal{D}^- = I - H$ . This yields

$$x_2 = Fq_1 - Fx'_1 - FB_{11}x_1 \quad (4.11)$$

$$(I - H)x'_1 + (I - H)B_{11}x_1 = (I - H)q_1, \quad (4.12)$$

where we used the identities  $FB_{12} = I$  and  $(I - H)B_{12} = 0$ .

Let  $u(t) = \mathcal{D}P_1x(t) = (I - H)x_1(t)$ . Then, we observe that

$$\begin{aligned} (I - H)B_{11}u &= (I - H)B_{11}(I - H)x_1 = (I - H)B_{11}x_1 - (I - H)B_{11}Hx_1, \text{ or} \\ (I - H)B_{11}x_1 &= (I - H)B_{11}u + (I - H)B_{11}Hx_1. \end{aligned}$$

Substituting the last line above into (4.12) along with  $u' = (I - H)x'_1$  yields

$$\begin{aligned} u' + (I - H)B_{11}u &= (I - H)q_1 - (I - H)B_{11}Hx_1 \\ u' + (I - H)B_{11}u &= (I - H)q_1 - (I - H)B_{11}Tq_2, \end{aligned} \quad (4.13)$$

where we also used  $Hx_1 = Tq_2$  from equation (4.10). Equation (4.13) is called the abstract ODE, or AODE, for this problem. We note that this AODE exists in the space  $\mathcal{H}_1$ , i.e.,  $u(t) \in \mathcal{H}_1$ . We therefore need an associated initial condition written as  $u(0) = \mathcal{D}P_1x(0) = (I - H)x_1(0) \in \mathcal{H}_1$  (without loss of generality we assume our initial time  $t_0 = 0$ ).

Since  $(I - H)^2 = (I - H)$  and  $\mathcal{B}$  are bounded, we can write the AODE (4.13) in the form

$$u'(t) + \mathcal{L}u(t) = \tilde{q}(t),$$

where  $\mathcal{L} = (I - H)B_{11}$  with  $\mathcal{L} \in L(\mathcal{Z})$ ,  $\mathcal{Z} = \mathcal{H}_1$ , and  $\tilde{q}(t)$  equals the right-hand side of the AODE. We now invoke Lemma 3.5 to give us existence and uniqueness of a solution  $u(t) \in C^1([0, T]; \mathcal{H}_1)$  to this AODE (4.13), where  $q \in C(0, T; \mathcal{Y})$  implies  $\tilde{q} \in C(0, T; \mathcal{Y})$ .

However, we need to add a stronger condition on  $q$ . Since we need the derivative with respect to time of  $u = (I - H)x_1 = x_1 - Hx_1$ ,  $Hx_1$  needs to be differentiable. If the left-hand side of (4.10) is differentiable, then so must the right-hand side be. Thus, we need  $q_2 \in C^1(0, T; \mathcal{H}_2)$  whereas  $q_1 \in C(0, T; \mathcal{H}_1)$ .

We are given a consistent Hessenberg initial condition for  $x_1(0)$ . Surprisingly, no initial condition on  $x_2(t)$  needs to be provided in order to solve for  $x(t)$ .  $x_2(0)$  will be determined by  $q_1(0)$ ,  $x_1(0)$ , and  $x'_1(0)$  and equation (4.11). Thus, the abstract Hessenberg DAE (4.1) is solvable for consistent Hessenberg initial condition  $x_1(0)$  that satisfies  $B_{21}x_1(0) = q_2(0)$ .

To obtain a solution representation, we start with  $u = (I - H)x_1 = x_1 - Hx_1$ . This gives  $x_1 = u + Hx_1$ . But from equation (4.10) we have  $Hx_1 = Tq_2$ . Thus,  $x_1 = u + Tq_2$ .

Since we have a unique  $u$  from the AODE, we will have a unique solution  $x_1$ . We note that both the AODE (4.13) and the preceding formula for  $x_1$  only depend on elements of the operator  $\mathcal{B}$ , the right-hand side forcing function  $q$  and the consistent initial condition  $x(0)$ , all of which are given data of the problem. In other words, they are not dependent on our particular choice of  $Q_0$ . Additionally, we have that  $x_1 \in C^1([0, T]; \mathcal{H}_1)$  since both  $u$  and  $q_2$  are continuously differentiable. Hence, once we have obtained a unique solution for  $x_1$ , we can obtain a unique solution  $x_2$  from equation (4.11). Thus, the solution may be represented as

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix},$$

where

$$\begin{aligned} x_1(t) &= u(t) + Tq_2(t) \\ x_2(t) &= Fq_1(t) - Fx'_1(t) - FB_{11}x_1(t) \end{aligned} \quad (4.14)$$

Substituting back into (4.8) and (4.9), it is easy to show that our solution representation for  $x_1(t)$  and  $x_2(t)$  from (4.14) satisfies the abstract Hessenberg DAE (4.1).

We can rewrite the solution representation (4.14) in terms of the unique solution  $u(t)$  of the AODE and the right-hand side function  $q(t)$  to yield

$$\begin{aligned} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} &= \begin{bmatrix} u(t) + Tq_2(t) \\ Fq_1(t) - Fx'_1(t) - FB_{11}x_1(t) \end{bmatrix} \\ &= \begin{bmatrix} u(t) + Tq_2(t) \\ Fq_1(t) - Fu'(t) - FTq'_2(t) - FB_{11}u(t) - FB_{11}Tq_2(t) \end{bmatrix} \\ &= \begin{bmatrix} u(t) + Tq_2(t) \\ Fq_1(t) - F[-(I - H)B_{11}u(t) + (I - H)q_1(t) - (I - H)B_{11}Tq_2(t)] \dots \\ -FTq'_2(t) - FB_{11}u(t) - FB_{11}Tq_2(t) \end{bmatrix} \\ &= \begin{bmatrix} u(t) + Tq_2(t) \\ -FB_{11}u(t) + Fq_1(t) - FB_{11}Tq_2(t) - (B_{21}B_{12})^{-1}q'_2(t) \end{bmatrix} \\ &= \begin{bmatrix} I \\ -FB_{11} \end{bmatrix} u(t) + \begin{bmatrix} Tq_2(t) \\ Fq_1(t) - FB_{11}Tq_2(t) - (B_{21}B_{12})^{-1}q'_2(t) \end{bmatrix}. \end{aligned} \quad (4.15)$$

where we used the identity  $F(I - H) = 0$  in some of the calculations. We note, it is obvious from this version of the solution representation that we need  $q_2 \in C^1(0, T; \mathcal{H}_2)$ .  $\square$

*Remark.* In the index-2 case, abstract DAEs can have “hidden” constraints. However, in the abstract Hessenberg DAE case (4.1) this is not the situation as the algebraic constraint equation is explicit in (4.9).

### 4.3 A Case Where $\mathcal{B}$ is Unbounded

We now weaken the assumption that  $\mathcal{B}$  is bounded. Specifically, we will look at the case of  $\mathcal{B}_{11} : \mathcal{D}(\mathcal{B}_{11}) \subset \mathcal{H}_1 \rightarrow \mathcal{H}_1$  unbounded with  $\mathcal{D}(\mathcal{B}_{11})$  dense in  $\mathcal{H}_1$ . We will leave  $B_{12}$  and  $B_{21}$  bounded with the auxiliary condition that  $B_{21}B_{12}$  is one-to-one and onto. In order to make this work we will have to impose an additional condition on  $\mathcal{D}(\mathcal{B}_{11})$  relative to  $\text{im } B_{12}$ . We can then state the following result.

**Theorem 4.2.** *Given the following conditions: (i)  $B_{12}$  and  $B_{21}$  bounded with the composition  $B_{21}B_{12}$  bijective, (ii)  $\mathcal{B}_{11}$  is a closed operator with dense domain  $\mathcal{D}(\mathcal{B}_{11}) \subset \mathcal{H}_1$  such that  $\text{im } B_{12} \subset \mathcal{D}(\mathcal{B}_{11})$ , (iii)  $\mathcal{B}_{11}$  is the infinitesimal generator of a strongly continuous semigroup on  $\mathcal{H}_1$ , (iv) the right-hand side function  $q = [q_1 \ q_2]^T$  satisfies  $q_1 \in C(0, T; \mathcal{H}_1)$  and  $q_2 \in C^1(0, T; \mathcal{H}_2)$ , and (v) a consistent Hessenberg initial condition  $x_1(0)$ .*

*Then there exists a unique solution  $x \in C([0, T]; \mathcal{X})$  to the semi-explicit abstract Hessenberg DAE (4.1).*

*Proof.* We first remark that since  $\text{im } B_{12} \subset \mathcal{D}(\mathcal{B}_{11})$ , the compositions  $\mathcal{B}_{11}T$  and  $\mathcal{B}_{11}H$  make sense where  $T$  and  $H$  are defined as in Section 4.1. Then the expressions for the canonical  $G_2$  and  $G_2^{-1}$  given in (4.6) and (4.7) are still valid. Furthermore, both remain bounded. Hence,  $G_2$  is bijective and this DAE remains index-2.

We now will work toward decoupling the abstract Hessenberg DAE into its AODE and its constraint equation in a manner similar to that in Section 4.2. We start with the two equations from the Hessenberg DAE (4.1):

$$x_1'(t) + B_{11}x_1(t) + B_{12}x_2(t) = q_1(t) \quad (4.16)$$

$$B_{21}x_1(t) = q_2(t), \quad (4.17)$$

where we assume  $x_1(t) \in \mathcal{D}(\mathcal{B}_{11})$ . As before we multiply the second equation (4.17) by  $T$ , a reflexive generalized inverse of  $B_{21}$ , to obtain

$$Hx_1 = Tq_2. \quad (4.18)$$

We then multiply the first equation (4.16) once by  $F$ , a reflexive generalized inverse of  $B_{12}$ , and again by the projector  $(I - H)$  to yield two equations:

$$x_2 = Fq_1 - Fx_1' - FB_{11}x_1, \quad (4.19)$$

$$(I - H)x_1' + (I - H)B_{11}x_1 = (I - H)q_1. \quad (4.20)$$

We let  $u(t) = (I - H)x_1(t)$  and derive the AODE as previously,

$$u' + (I - H)B_{11}u = (I - H)q_1 - (I - H)B_{11}Tq_2. \quad (4.21)$$

We note that  $u \in \mathcal{D}(\mathcal{B}_{11}) \subset \mathcal{H}_1$  so this equation is valid.

We will next utilize the following result on perturbation of a semigroup generator from Engel and Nagel [12].

**Lemma 4.1.** *Given a Banach space  $\mathcal{X}$  with  $\mathcal{D}(\mathcal{A}) \subset \mathcal{X}$ . Define  $\mathcal{X}_1^{\mathcal{A}} = \mathcal{D}(\mathcal{A})$  with the graph norm  $\|x\|_1 := \|x\|_{\mathcal{X}} + \|\mathcal{A}x\|_{\mathcal{X}}$ . If  $B \in L(\mathcal{X}_1^{\mathcal{A}}, \mathcal{X})$  and  $\mathcal{A}$  is an infinitesimal generator of a strongly continuous semigroup on  $\mathcal{X}$ , then  $\mathcal{A} + B$  with  $\mathcal{D}(\mathcal{A} + B) = \mathcal{D}(\mathcal{A})$  is an infinitesimal generator of a strongly continuous semigroup on  $\mathcal{X}$ .*

We have  $(I - H)\mathcal{B}_{11} = \mathcal{B}_{11} - H\mathcal{B}_{11}$ . We then need to show that  $-H\mathcal{B}_{11}$  is bounded on  $\mathcal{X}_1^{\mathcal{B}_{11}} := (\mathcal{D}(\mathcal{B}_{11}), \|\cdot\|_1)$ . We now restrict the operators  $H$  and  $\mathcal{B}_{11}$  to  $\mathcal{D}(\mathcal{B}_{11})$ . We then have

$$\begin{aligned} \|-H\mathcal{B}_{11}x\|_1 &= \|H\mathcal{B}_{11}x\| + \|\mathcal{B}_{11}H\mathcal{B}_{11}x\| \\ &\leq \|H\|\|\mathcal{B}_{11}x\| + \|\mathcal{B}_{11}H\|\|\mathcal{B}_{11}x\| \\ &\leq C_1\|\mathcal{B}_{11}x\| + C_2\|\mathcal{B}_{11}x\| \\ &= C\|\mathcal{B}_{11}x\| \\ &\leq C\|x\| + C\|\mathcal{B}_{11}x\| \\ &= C\|x\|_1, \end{aligned}$$

where we used Proposition 3.1, part (viii), that  $\mathcal{B}_{11}H$  is a bounded operator by the Closed Graph Theorem since it is defined on the entire space  $\mathcal{X}_1 = \mathcal{D}(\mathcal{B}_{11})$ . Thus,  $-H\mathcal{B}_{11} \in L(\mathcal{X}_1^{\mathcal{B}_{11}}, \mathcal{H}_1)$ . Then, by Lemma 4.1,  $(I - H)\mathcal{B}_{11}$  is an infinitesimal generator of a strongly continuous semigroup on  $\mathcal{H}_1$ .

Since we have a consistent Hessenberg initial condition  $x_1(0)$ , we then form  $u(0) = (I - H)x_1(0)$ . Then by invoking Lemma 3.5, we can again obtain existence and uniqueness of a solution  $u \in C^1([0, T]; \mathcal{H}_1)$  to the AODE (4.21). The solution to the DAE can then be found to be the unique representation

$$\begin{aligned} x_1(t) &= u(t) + Tq_2(t) \\ x_2(t) &= Fq_1(t) - Fx_1'(t) - FB_{11}x_1(t). \end{aligned}$$

As above,  $x_1(t)$  is continuously differentiable. Similarly, we can also write this representation in the form (4.15). As an end result, we also have that  $x \in C([0, T]; \mathcal{X})$ .  $\square$

*Remark.* It is possible to revise some of the hypotheses of Theorem 4.2. If we have that  $\mathcal{B}_{11}$  is the infinitesimal generator of a strongly continuous semigroup of contractions on  $\mathcal{H}_1$  and if the composition  $-H\mathcal{B}_{11}$  is dissipative, then one could use instead Lemma 3.8 above to get a unique solution to the AODE. One would then proceed as before to obtain a solution representation.

*Remark.* As seen in applications in the next chapter, it is not uncommon for the space  $\mathcal{H}_1$  to be a product space such as  $\mathcal{X}_1 \times \mathcal{X}_2$ . Let  $\mathcal{Y}_2$  be dense in  $\mathcal{X}_2$ . Then, if  $\text{im } B_{12} = \{0\} \times \mathcal{Y}_2$  with  $\mathcal{D}(\mathcal{B}_{11}) = \mathcal{X}_1 \times \mathcal{Y}_2$  we will have  $\text{im } B_{12} \subset \mathcal{D}(\mathcal{B}_{11})$ . Thus, our assumption may not necessarily be that restrictive.

We may also be able to weaken the condition  $\text{im } B_{12} \subset \mathcal{D}(\mathcal{B}_{11})$ . Let us define the concept of *uniformly dense with respect to  $P$*  as follows:

**Definition 4.2.** If  $P$  is an orthogonal projector on a Hilbert space  $\mathcal{H}$  and the domain of an operator  $\mathcal{A}$ ,  $\mathcal{D}(\mathcal{A})$ , is dense in  $\mathcal{H}$  such that  $\mathcal{D}(\mathcal{A}) \cap \text{im } P$  is also dense in  $\text{im } P$  and  $\mathcal{D}(\mathcal{A}) \cap \text{im } (I - P)$  is also dense in  $\text{im } (I - P)$ , then we say  $\mathcal{D}(\mathcal{A})$  is *uniformly dense with respect to  $P$* .

We can then apply the following Lemma.

**Lemma 4.2.** *Let the domain of an operator  $\mathcal{A}$ ,  $\mathcal{D}(\mathcal{A})$ , be dense in a Hilbert space  $\mathcal{H}$ . Let  $P$  be an orthogonal projector in  $\mathcal{H}$ . Furthermore, let  $\mathcal{D}(\mathcal{A})$  be uniformly dense with respect to  $P$ . Then,  $\mathcal{D}(\mathcal{A})$  restricted to those elements  $h \in \mathcal{D}(\mathcal{A}) \subset \mathcal{H}$  such that  $Ph \in \mathcal{D}(\mathcal{A})$  is also dense in  $\mathcal{H}$ , i.e.,  $\mathcal{D}(\mathcal{A})|_P := \{h \in \mathcal{D}(\mathcal{A}) \subset \mathcal{H} : Ph \in \mathcal{D}(\mathcal{A})\}$ .*

*Proof.* Let  $h \in \mathcal{H}$  be arbitrary. Then, since  $\mathcal{D}(\mathcal{A})$  is dense in  $\mathcal{H}$ , there is an element  $x \in \mathcal{D}(\mathcal{A})$  such that given arbitrary  $\epsilon > 0$ ,  $\|h - x\|_{\mathcal{H}} < \frac{\epsilon}{3}$ . Since  $P$  is an orthogonal projector on  $\mathcal{H}$ , we can write  $x = y + z$  where  $y = Px$  and  $z = (I - P)x$ . If  $y \in \mathcal{D}(\mathcal{A})$  we are done since then  $x \in \mathcal{D}(\mathcal{A})|_P$ .

If  $y \notin \mathcal{D}(\mathcal{A})$  we proceed as follows. Since  $\text{im } P$  for a projector in a Hilbert space is a closed linear subspace and  $\mathcal{D}(\mathcal{A})$  is dense in  $\mathcal{H}$ , by  $\mathcal{D}(\mathcal{A})$  uniformly dense with respect to  $P$  there exists a  $\tilde{y} \in \mathcal{D}(\mathcal{A}) \cap \text{im } P$  such that  $\|\tilde{y} - y\|_{\mathcal{H}} < \frac{\epsilon}{3}$ . Define  $\tilde{x} := \tilde{y} + z$ . Then, by definition,  $\|\tilde{x} - x\|_{\mathcal{H}} = \|\tilde{y} + z - y - z\|_{\mathcal{H}} = \|\tilde{y} - y\|_{\mathcal{H}} < \frac{\epsilon}{3}$ . If  $\tilde{x} \in \mathcal{D}(\mathcal{A})|_P$ , then we are done since  $\|h - \tilde{x}\|_{\mathcal{H}} = \|h - x\|_{\mathcal{H}} + \|x - \tilde{x}\|_{\mathcal{H}} < \frac{\epsilon}{3} + \frac{\epsilon}{3} < \epsilon$ .

However, if  $\tilde{x} \notin \mathcal{D}(\mathcal{A})|_P$ , similarly by  $\mathcal{D}(\mathcal{A})$  uniformly dense with respect to  $P$ , there is an element  $\hat{x} = \tilde{y} + \hat{z}$  such that  $\hat{x} \in \mathcal{D}(\mathcal{A})$  and  $\|\hat{x} - \tilde{x}\|_{\mathcal{H}} < \frac{\epsilon}{3}$ . Note that we only change  $\tilde{x}$  to  $\hat{x}$  by perturbing  $z$  within  $\text{im } (I - P)$  which is also a linear closed subspace of  $\mathcal{H}$ . Furthermore, by definition,  $\hat{x} \in \mathcal{D}(\mathcal{A})|_P$ . Finally,  $\|h - \hat{x}\|_{\mathcal{H}} = \|h - x\|_{\mathcal{H}} + \|x - \tilde{x}\|_{\mathcal{H}} + \|\tilde{x} - \hat{x}\|_{\mathcal{H}} < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$ . Thus,  $\mathcal{D}(\mathcal{A})|_P$  is also dense in  $\mathcal{H}$ .  $\square$

*Remark.* We can reapply this Lemma again if desired. For example, assume  $\mathcal{D}(\mathcal{A}) = \mathcal{D}(\mathcal{A})|_P$ , we can then restrict this new domain to  $\mathcal{D}(\mathcal{A})|_{I-P}$ . We then have that the restricted domain is also dense. We thus end up with the domain defined as  $\{h \in \mathcal{D}(\mathcal{A}) \subset \mathcal{H} : Ph \in \mathcal{D}(\mathcal{A}) \text{ and } (I - P)h \in \mathcal{D}(\mathcal{A})\}$  is also dense in  $\mathcal{H}$ .

Thus, we could require that  $\mathcal{D}(\mathcal{B}_{11})$  be uniformly dense with respect to  $H$ , apply the Lemma, and then restrict  $u(t)$  in the AODE to  $\mathcal{D}(\mathcal{B}_{11})|_H$  which would then be dense in  $\mathcal{H}_1$ .

# Chapter 5

## Applications

We start with some applications involving systems of partial differential equations. Both applications lead to index-1 abstract DAEs. The first example arises from a thermal model of a chemical vapor deposition reactor for growing superconducting films. The second example is much more detailed and includes the complete derivation and modeling of a structural dynamics problem as a DAE.

### 5.1 Thermal Model for Chemical Vapor Deposition Reactor

This application is related to the processing of superconducting films for use in high performance microwave filters for cellular communications. One method involves growing these films in a low-pressure chemical vapor deposition reactor. In this process chemical reactants are sprayed onto a heated substrate. The goal is to then grow a thin uniform film on this substrate. It is expected that conduction and radiation will be the dominant thermal mechanisms vice convection since the sprayed fluid has a low thermal capacitance. The thermal controls used in this model will be the power supplied to the individual heater rings which heat the substrate disk and the temperature of the reactor walls which is controlled via an oil-cooled heat exchanger. A diagram of the key items needed for the thermal model is shown in Figure 5.1. For ease of modeling, the substrate as shown in this figure is actually upside down, that is, the thin film will be grown on the bottom of the disk shown in the figure.

Cliff and Herdman [9] developed the model that we use here. For the disk, conduction



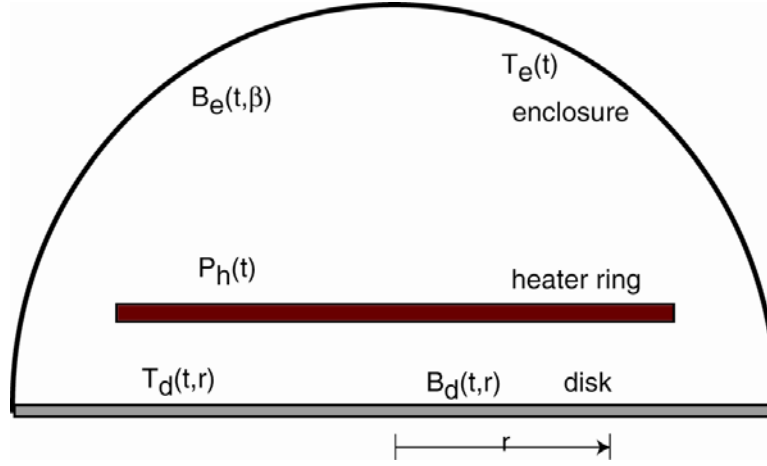


Figure 5.1: Simplified Thermal Model of a Chemical Vapor Deposition Reactor

in a thin, circular plate is modeled by a diffusion equation:

$$\frac{\partial T_d(t, r)}{\partial t} = \frac{\kappa}{\rho C_p r} \left[ \frac{\partial}{\partial r} \left( r \frac{\partial T_d(t, r)}{\partial r} \right) \right] - \frac{S(t, r)}{\rho C_p \tau}, \quad 0 < r < R, \quad t > 0, \quad (5.1)$$

where  $T_d(t, r)$  is the surface temperature of the disk at distance  $r$  from the center. We assume radial symmetry of the temperature distribution on the disk.  $S(t, r)$  is the net heat flux from the surface,  $\tau$  is the disk thickness which is assumed to be small compared to its radius  $R$ ,  $\rho$  is its density,  $C_p$  is the specific heat of the disk, and  $\kappa$  is the thermal conductivity. Thus,  $\frac{\kappa}{\rho C_p}$  is the usual thermal diffusivity of the disk. We also specify an initial temperature distribution on the disk  $T_d(0, r) = g(r)$ .

The source term involving the net heat flux arises from radiant exchange among the heater, the disk and the enclosure. It can be expressed as

$$S(t, r) = \frac{\epsilon}{1 - \epsilon} (\sigma T^4(t, r) - B_d(t, r)), \quad (5.2)$$

where  $B_d(t, r)$  is the disk radiosity or radiant energy flux departing from the disk. We then have a coupled integral equation involving the disk radiosity  $B_d(t, r)$ , the enclosure radiosity  $B_e(t, r)$  and the disk temperature  $T_d(t, r)$ ,

$$(I - F) \begin{bmatrix} B_d(t, r) \\ B_e(t, r) \end{bmatrix} = \begin{bmatrix} \epsilon_d \sigma T_d^4(t, r) \\ 0 \end{bmatrix} + P \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}, \quad (5.3)$$

where  $u_1(t)$  is the heater power and  $u_2(t)$  is the fourth power of the enclosure uniform surface temperature  $T_e(t)$ .  $F$  is a linear Fredholm integral operator such that  $\|F\| < 1$  and  $P$  is a bounded linear operator.

We now put the system into abstract equation form. Let  $x_1(t) = T_d(t, \cdot)$  and

$$x_2(t) := \begin{bmatrix} B_d(t, \cdot) \\ B_e(t, \cdot) \end{bmatrix}.$$

Then, we can write equation (5.1) as

$$\begin{aligned} x_1'(t) + Kx_1(t) + \mathcal{F}(x_1(t), x_2(t)) &= 0, \\ \mathcal{N}(x_1(t)) + (I - F)x_2(t) &= q_2(t), \end{aligned}$$

where prime indicates differentiation with respect to time,  $\mathcal{F}$  is a nonlinear operator involving the fourth power of  $x_1(t)$ ,  $\mathcal{N}$  is also a nonlinear operator involving the fourth power of  $x_1(t)$ , and  $q_2(t) = P[u_1(t) \ u_2(t)]^T$ . We will assume for simplicity that we can linearize both  $\mathcal{F}$  and  $\mathcal{N}$ . We then incorporate the linearized term for  $x_1(t)$  into the operator  $K$  and call the remaining linear operator  $L$ . We label the linearized operator for  $\mathcal{N}$  as  $E$ . This yields the following *linearized* equations

$$\begin{aligned} x_1'(t) + Kx_1(t) + Lx_2(t) &= 0, \\ Ex_1(t) + (I - F)x_2(t) &= q_2(t), \end{aligned}$$

where  $L$  and  $E$  are linear bounded operators and  $Kz = -\Delta z + cz$  with  $c > 0$ . We then put this in semi-explicit DAE form (3.2)

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} + \begin{bmatrix} K & L \\ E & (I - F) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 0 \\ q_2(t) \end{bmatrix}. \quad (5.4)$$

Since  $\|F\| < 1$  we have that  $(I - F)$  is one-to-one and onto by a standard functional analysis theorem (for example, see Theorem III.1.4 in Taylor and Lay [38]). Furthermore,  $E$  and  $L$  are bounded linear operators. Curtain and Pritchard [10], as well as Goldstein [17] establish for the operator  $K$  with domain dense in a Hilbert space that  $K$  is the infinitesimal generator of an analytic semigroup.

Finally, by Theorem 3.2 with  $q_2(t) \in C^1([0, T]; \mathcal{H})$  and with consistent index-1 initial condition, there exists a unique solution to the semi-explicit DAE (5.4) for this linearized thermal model of a chemical vapor deposition reactor.

*Remark.* In this example, previous work was done by Cliff and Herdman [9] to put the model into DAE form. We did not go into the details of that construction. In contrast, the next application will show how a DAE model is developed and implemented.

## 5.2 Coupled Transversal Motion of Two Beams Connected by a Rigid Joint

The Defense Advanced Research Projects Agency (DARPA) is investigating stability and dynamics of large space arrays. These arrays involve large trusses made up of pyramidal

structures of beams. Figure 5.2 shows part of such a truss array. To simplify the analyses, we look at a simple model involving just two beams, each fixed at one end and then connected together at the other end by a rigid joint. See Figure 5.3 for a diagram. While the joint restricts movement, bending and vibrations can and do occur. Hence, the motivation to analyze such structures for stability and controllability.

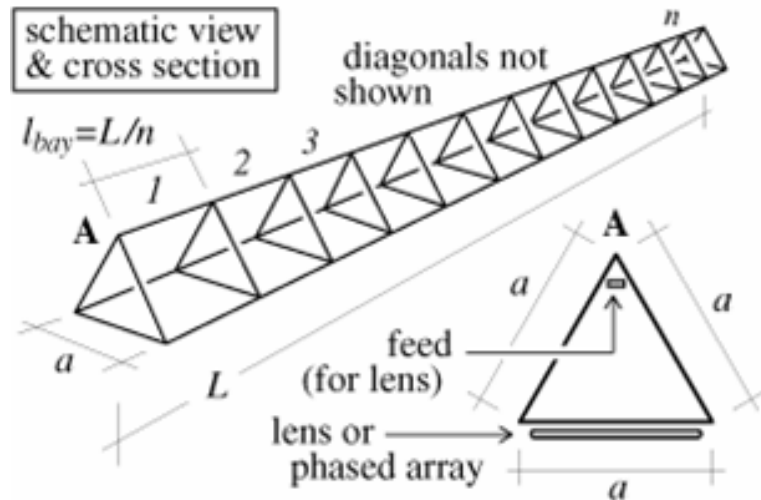


Figure 5.2: An Example of a Truss Array

For our investigation, we model the problem in an abstract DAE format and show existence and uniqueness of solutions under certain conditions.

### 5.2.1 Modeling and Development of Abstract DAE

We model the bending motions of the two identical beams using the Euler-Bernoulli beam partial differential equations with Kelvin-Voigt damping. We include homogeneous boundary conditions at the fixed ends of the beams. For simplicity we assume each beam is identical in this space truss. This leads to one partial differential equation for each

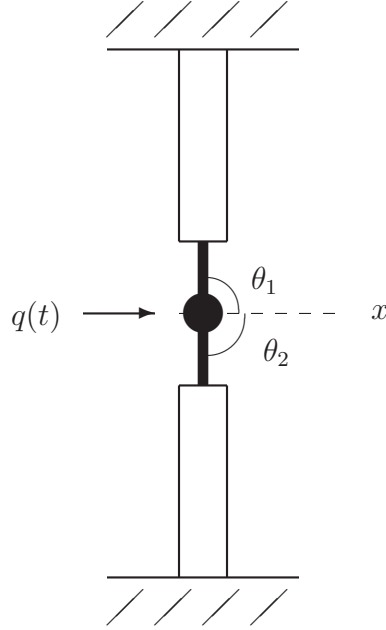


Figure 5.3: Model of Two Beams and a Joint

beam as follows (subscript 1 refers to beam 1 and subscript 2 refers to beam 2):

$$\begin{aligned}
 \rho A \frac{\partial^2 w_1(t, s_1)}{\partial t^2} + \frac{\partial^2}{\partial s_1^2} \left[ EI \frac{\partial^2 w_1(t, s_1)}{\partial s_1^2} + \gamma \frac{\partial^3 w_1(t, s_1)}{\partial s_1^2 \partial t} \right] &= 0 \quad (\text{beam 1}) \\
 w_1(t, 0) = \frac{\partial w_1(t, 0)}{\partial s_1} &= 0 \\
 \rho A \frac{\partial^2 w_2(t, s_2)}{\partial t^2} + \frac{\partial^2}{\partial s_2^2} \left[ EI \frac{\partial^2 w_2(t, s_2)}{\partial s_2^2} + \gamma \frac{\partial^3 w_2(t, s_2)}{\partial s_2^2 \partial t} \right] &= 0 \quad (\text{beam 2}) \\
 w_2(t, 0) = \frac{\partial w_2(t, 0)}{\partial s_2} &= 0,
 \end{aligned} \tag{5.5}$$

where  $w_i(t, s_i)$  represents the transverse displacement from the  $y$ -axis along beam  $i$  at time  $t$ ,  $s_i$  is the distance along beam  $i$  as measured from the fixed end 0 to the end connected to the joint where  $s_i = L$ , the length of each beam. Since the beams are identical the physical parameters  $\rho$ ,  $A$ ,  $E$ , and  $\gamma$  are the same for each beam, where  $A$  is the cross-sectional area,  $\rho$  is the mass density,  $E$  is the Young's modulus, and  $\gamma$  is the Kelvin-Voigt damping parameter for each beam. We will also assume an appropriate initial configuration of the beams at  $t = 0$ , i.e.,  $w_1(0, s_1) = f_1(s_1)$  and  $w_2(0, s_2) = f_2(s_2)$ .

The joint consists of a pivot point with two legs. Each leg is attached to the end of a

beam. Once again the two joint legs will be considered identical with  $m$  the mass of each joint leg,  $m_p$  the mass of the pivot,  $M = 2m + m_p$  the total mass of the entire joint structure,  $\ell$  the length of each joint leg,  $d$  the distance from the center of mass of each joint leg to the pivot point,  $I$  the moment of inertia of each joint leg about its center of mass.

For simplicity of analysis, we further assume the two beams and joint start out close to a straight vertical alignment and that the amplitude of vibrations are small compared to the physical dimensions of the system. We therefore use linearized equations of motion derived from Newton's Second Law of Motion. This leads to a system of ordinary differential equations describing the motion of the joint pivot point and angles for the joint legs. We define the vector-valued function  $y(t) = (x(t), \theta_1(t), \theta_2(t))^T \in \mathbb{R}^3$  where  $x(t)$  represents the horizontal motion of the joint pivot along the  $x$ -axis (for simplicity we ignore motion in the  $y$ -axis direction),  $\theta_1(t)$  represents the deviation angle of joint leg 1 from the vertical and  $\theta_2(t)$  is for joint leg 2. We further assume that the joint has internal viscous damping but no stiffness component. This leads to the system of ordinary differential equations describing the motion of the joint pivot and its legs:

$$\begin{aligned} \mathbf{M}y''(t) + \mathbf{W}y'(t) - \mathbf{L}F(t) &= q(t), \text{ or} \\ y''(t) + \mathbf{M}^{-1}\mathbf{W}y'(t) - \mathbf{M}^{-1}\mathbf{L}F(t) &= \mathbf{M}^{-1}q(t), \end{aligned} \quad (5.6)$$

where the prime ( $\prime$ ) indicates differentiation with respect to time  $t$  and

$$\mathbf{M} = \begin{bmatrix} M & -md & md \\ -md & I + md^2 & 0 \\ md & 0 & I + md^2 \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & b & -b \\ 0 & -b & b \end{bmatrix},$$

with  $b \geq 0$  the damping coefficient. We remark at this point that  $\mathbf{M}$  is positive definite symmetric. We also have

$$\mathbf{L} = \begin{bmatrix} 0 & -1 & 0 & 1 \\ 1 & \ell & 0 & 0 \\ 0 & 0 & 1 & \ell \end{bmatrix},$$

where  $\mathbf{L}$  has full row rank.

For some generality, we assume an external forcing function  $q(t) = [f(t) \ 0 \ 0]^T \in \mathbb{R}^3$  is applied at the pivot point. We define  $F(t) = (M_1(t), N_1(t), M_2(t), N_2(t))^T \in \mathbb{R}^4$  as the applied load on the joint resulting from motion of the two attached beams, where  $M_i(t)$  is the effective torque applied to the joint pivot point from beam  $i$  and  $N_i(t)$  is the effective shear load applied to the center of mass of joint leg  $i$  due to beam  $i$ . These loads can be represented in terms of the  $w_i(t, s_i)$  variables from the beam partial differential

equations. The resulting boundary compatibility constraints are:

$$F(t) = \begin{bmatrix} M_1(t) \\ N_1(t) \\ M_2(t) \\ N_2(t) \end{bmatrix} = \begin{bmatrix} EI \frac{\partial^2 w_1(t, L)}{\partial s_1^2} + \gamma \frac{\partial^3 w_1(t, L)}{\partial s_1^2 \partial t} \\ EI \frac{\partial^3 w_1(t, L)}{\partial s_1^3} + \gamma \frac{\partial^4 w_1(t, L)}{\partial s_1^3 \partial t} \\ EI \frac{\partial^2 w_2(t, L)}{\partial s_2^2} + \gamma \frac{\partial^3 w_2(t, L)}{\partial s_2^2 \partial t} \\ EI \frac{\partial^3 w_2(t, L)}{\partial s_2^3} + \gamma \frac{\partial^4 w_2(t, L)}{\partial s_2^3 \partial t} \end{bmatrix}. \quad (5.7)$$

As for the partial differential equations for the beams, we also assume an appropriate initial condition for the configuration of the joint pivot location and joint leg angles, i.e.,  $y(0) = y_0 \in \mathbb{R}^3$ .

We next have some geometric compatibility constraints that realize the connection between the ends of the joint leg and the ends of the beams. These conditions require that the ends of the joint legs and beams have the same physical coordinates and that the end slope of the beams remain aligned with that of the joint legs. This leads to the following constraint equation:

$$\begin{bmatrix} \frac{\partial w_1(t, L)}{\partial s_1} \\ w_1(t, L) \\ \frac{\partial w_2(t, L)}{\partial s_2} \\ w_2(t, L) \end{bmatrix} = \mathbf{L}^T y(t). \quad (5.8)$$

We now have all the governing equations describing the two beams and joint system.

We let

$$w(t) = \begin{bmatrix} w_1(t, \cdot) \\ w_2(t, \cdot) \end{bmatrix} \in \mathcal{H}_1 = \{H_\ell^2(0, L) \cap H^4(0, L)\} \times \{H_\ell^2(0, L) \cap H^4(0, L)\},$$

where  $H_\ell^2(0, L) = \{w_i(t, \cdot) \in H^2(0, L) : w_i(t, 0) = \frac{\partial w_i(t, 0)}{\partial s_i} = 0\}$ .

We define  $w'(t)$ , the derivative of  $w(t)$  with respect to time  $t$ , similarly with  $w'(t) \in \mathcal{H}_2 = \{H_\ell^2(0, L) \cap H^4(0, L)\} \times \{H_\ell^2(0, L) \cap H^4(0, L)\}$ .

We then define

$$\sigma(t) = \begin{bmatrix} \sigma_1(t) \\ \sigma_2(t) \end{bmatrix} = \begin{bmatrix} EI \frac{\partial^2 w_1(t, \cdot)}{\partial s_1^2} + \gamma \frac{\partial^3 w_1(t, \cdot)}{\partial s_1^2 \partial t} \\ EI \frac{\partial^2 w_2(t, \cdot)}{\partial s_2^2} + \gamma \frac{\partial^3 w_2(t, \cdot)}{\partial s_2^2 \partial t} \end{bmatrix}, \text{ or}$$

$$\sigma(t) = \mathbf{E}D^2w(t) + \mathbf{G}D^2w'(t), \quad (5.9)$$

with  $\sigma(t) \in \mathcal{H}_\sigma = H^2(0, L) \times H^2(0, L)$  and

$$\mathbf{E} = \begin{bmatrix} EI & 0 \\ 0 & EI \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \gamma & 0 \\ 0 & \gamma \end{bmatrix}, \quad D^2 = \begin{bmatrix} \frac{\partial^2}{\partial s_1^2} & 0 \\ 0 & \frac{\partial^2}{\partial s_2^2} \end{bmatrix}.$$

We can now rewrite the beam partial differential equations (5.5) as an abstract second order ODE:

$$w''(t) + \mathbf{A}D^2\sigma(t) = 0, \quad (5.10)$$

where

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\rho A} & 0 \\ 0 & \frac{1}{\rho A} \end{bmatrix}.$$

We next define the boundary operator  $\mathcal{T} : \{H^2(0, L) \times H^2(0, L)\} \rightarrow \mathbb{R}^4$  as

$$\mathcal{T}w(t) = \begin{bmatrix} -\frac{\partial w_1(t, L)}{\partial s_1} \\ w_1(t, L) \\ \frac{\partial w_2(t, L)}{\partial s_2} \\ w_2(t, L) \end{bmatrix},$$

and rewrite the geometric compatibility constraints (5.8) in the form

$$\mathcal{T}w(t) = \mathbf{L}^T y(t). \quad (5.11)$$

If we differentiate (5.11) with respect to time we get

$$\mathcal{T}w'(t) = \mathbf{L}^T y'(t). \quad (5.12)$$

We use the boundary operator  $\mathcal{T}$  to also rewrite our boundary compatibility constraints (5.7). We first note that calculating  $\mathcal{T}\sigma(t)$  yields:

$$\begin{aligned}
\mathcal{T}\sigma(t) &= \mathcal{T}(\mathbf{E}D^2w(t) + \mathbf{G}D^2w'(t)) \\
&= \mathcal{T}\mathbf{E}D^2w(t) + \mathcal{T}\mathbf{G}D^2w'(t) \\
&= \mathcal{T} \begin{bmatrix} EI \frac{\partial^2 w_1(t, \cdot)}{\partial s_1^2} \\ EI \frac{\partial^2 w_2(t, \cdot)}{\partial s_2^2} \end{bmatrix} + \mathcal{T} \begin{bmatrix} \gamma \frac{\partial^3 w_1(t, \cdot)}{\partial s_1^2 \partial t} \\ \gamma \frac{\partial^3 w_2(t, \cdot)}{\partial s_2^2 \partial t} \end{bmatrix} \\
&= \begin{bmatrix} -EI \frac{\partial^3 w_1(t, L)}{\partial s_1^3} - \gamma \frac{\partial^4 w_1(t, L)}{\partial s_1^3 \partial t} \\ EI \frac{\partial^2 w_1(t, L)}{\partial s_1^2} + \gamma \frac{\partial^3 w_1(t, L)}{\partial s_1^2 \partial t} \\ -EI \frac{\partial^3 w_2(t, L)}{\partial s_2^3} - \gamma \frac{\partial^4 w_2(t, L)}{\partial s_2^3 \partial t} \\ EI \frac{\partial^2 w_2(t, L)}{\partial s_2^2} + \gamma \frac{\partial^3 w_2(t, L)}{\partial s_2^2 \partial t} \end{bmatrix} = \begin{bmatrix} -N_1(t) \\ M_1(t) \\ -N_2(t) \\ M_2(t) \end{bmatrix}.
\end{aligned}$$

We then multiply both sides by a permutation or rearrangement matrix  $\mathbf{R}$ ,

$$\mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

to get

$$\begin{aligned}
\mathbf{R}\mathcal{T}\sigma(t) &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} -N_1(t) \\ M_1(t) \\ -N_2(t) \\ M_2(t) \end{bmatrix} = \begin{bmatrix} M_1(t) \\ N_1(t) \\ M_2(t) \\ N_2(t) \end{bmatrix} = F(t) \text{ or} \\
\mathbf{R}\mathcal{T}\sigma(t) &= F(t). \tag{5.13}
\end{aligned}$$

Thus, the boundary compatibility constraint can be represented by the operator constraint equation  $\mathbf{R}\mathcal{T}\sigma(t) = F(t)$ .

We now put this all together as a system of equations from (5.10), (5.6), (5.13), (5.9):

$$\begin{aligned}
w''(t) + \mathbf{A}D^2\sigma(t) &= 0 \\
y''(t) + \mathbf{M}^{-1}\mathbf{W}y'(t) - \mathbf{M}^{-1}\mathbf{L}F(t) &= \mathbf{M}^{-1}q(t) \\
\mathbf{R}\mathcal{T}\sigma(t) - F(t) &= 0 \\
\mathbf{E}D^2w(t) + \mathbf{G}D^2w'(t) - \sigma(t) &= 0.
\end{aligned}$$



Let  $z_1(t) = w(t)$ ,  $z_2(t) = z_1'(t) = w'(t)$ ,  $z_3(t) = y'(t)$ ,  $z_4(t) = \sigma(t)$ , and  $z_5(t) = F(t)$ . This leads to the following first order equations:

$$\begin{aligned} z_1'(t) - z_2(t) &= 0 \\ z_2'(t) + \mathbf{A}D^2z_4(t) &= 0 \\ z_3'(t) + \mathbf{M}^{-1}\mathbf{W}z_3(t) - \mathbf{M}^{-1}\mathbf{L}z_5(t) &= \mathbf{M}^{-1}q(t) \\ \mathbf{R}\mathcal{T}z_4(t) - z_5(t) &= 0 \\ \mathbf{E}D^2z_1(t) + \mathbf{G}D^2z_2(t) - z_4(t) &= 0. \end{aligned}$$

*Remark.* By Fattorini [14], we can reduce such a second order system to a first order system.

We then can put the above equations in matrix form:

$$\begin{bmatrix} I & 0 & 0 & \vdots & 0 & 0 \\ 0 & I & 0 & \vdots & 0 & 0 \\ 0 & 0 & I & \vdots & 0 & 0 \\ \dots & \dots & \dots & \vdots & \dots & \dots \\ 0 & 0 & 0 & \vdots & 0 & 0 \\ 0 & 0 & 0 & \vdots & 0 & 0 \end{bmatrix} z'(t) + \begin{bmatrix} 0 & -I & 0 & \vdots & 0 & 0 \\ 0 & 0 & 0 & \vdots & \mathbf{A}D^2 & 0 \\ 0 & 0 & \mathbf{M}^{-1}\mathbf{W} & \vdots & 0 & -\mathbf{M}^{-1}\mathbf{L} \\ \dots & \dots & \dots & \vdots & \dots & \dots \\ 0 & 0 & 0 & \vdots & \mathbf{R}\mathcal{T} & -I \\ \mathbf{E}D^2 & \mathbf{G}D^2 & 0 & \vdots & -I & 0 \end{bmatrix} z(t) = \begin{bmatrix} 0 \\ 0 \\ \mathbf{M}^{-1}q(t) \\ \dots \\ 0 \\ 0 \end{bmatrix}, \quad (5.14)$$

where  $z(t) = (z_1(t), z_2(t), z_3(t), z_4(t), z_5(t))^T$ . This equation is in the general form of  $\mathcal{E}z'(t) + \mathcal{B}z(t) = f(t)$  with  $\mathcal{E}$  a noninvertible operator. Furthermore, the domain of  $\mathcal{B}$ ,  $\mathcal{D}(\mathcal{B})$ , is defined as  $\mathcal{D}(\mathcal{B}) = \{z \in \mathcal{H}_1 \times \mathcal{H}_2 \times \mathbb{R}^3 \times \mathcal{H}_\sigma \times \mathbb{R}^4 : \mathcal{T}z_2 = \mathbf{L}^T z_3\}$  where we have incorporated the differentiated version of the geometric compatibility constraints (5.12) into the domain of our operator  $\mathcal{B}$ . Since the constraint  $\mathcal{T}z_2 = \mathbf{L}^T z_3$  only places conditions on  $z_2(t) = w'(t)$  on the right side (i.e., when  $s = L$ ) and not the left, we have that  $\mathcal{D}(\mathcal{B})$  is dense in  $\mathcal{H} = \{L^2(0, L) \times L^2(0, L)\} \times \{L^2(0, L) \times L^2(0, L)\} \times \mathbb{R}^3 \times \{L^2(0, L) \times L^2(0, L)\} \times \mathbb{R}^4$ .

We now partition the operator matrices as indicated by the dotted lines above to obtain

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} z'(t) + \begin{bmatrix} \mathcal{B}_{11} & \mathcal{B}_{12} \\ \mathcal{B}_{21} & \mathcal{B}_{22} \end{bmatrix} z(t) = \begin{bmatrix} q_1(t) \\ 0 \end{bmatrix}, \quad (5.15)$$

where

$$\mathcal{B}_{11} = \begin{bmatrix} 0 & -I & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathbf{M}^{-1}\mathbf{W} \end{bmatrix}, \quad \mathcal{B}_{21} = \begin{bmatrix} 0 & 0 & 0 \\ \mathbf{E}D^2 & \mathbf{G}D^2 & 0 \end{bmatrix},$$

and

$$\mathcal{B}_{12} = \begin{bmatrix} 0 & 0 \\ \mathbf{A}D^2 & 0 \\ 0 & -\mathbf{M}^{-1}\mathbf{L} \end{bmatrix}, \quad \mathcal{B}_{22} = \begin{bmatrix} \mathbf{R}\mathcal{T} & -I \\ -I & 0 \end{bmatrix}, \quad q_1(t) = \begin{bmatrix} 0 \\ 0 \\ \mathbf{M}^{-1}q(t) \end{bmatrix}.$$

We then have that  $\mathcal{D}(\mathcal{B}_{11}) = \mathcal{D}(\mathcal{B}_{21}) = \{(z_1, z_2, z_3)^T \in \mathcal{H}_1 \times \mathcal{H}_2 \times \mathbb{R}^3 : \mathcal{T}z_2 = \mathbf{L}^T z_3\}$  and  $\mathcal{D}(\mathcal{B}_{12}) = \mathcal{D}(\mathcal{B}_{22}) = \{(z_4, z_5)^T \in \mathcal{H}_\sigma \times \mathbb{R}^4\}$ .

We first investigate  $\mathcal{B}_{22}$  where  $\mathcal{B}_{22} : \mathcal{D}(\mathcal{B}_{22}) \rightarrow \mathcal{Y}_2 = \mathbb{R}^4 \times \{L^2(0, L) \times L^2(0, L)\}$ . We see that  $\ker \mathcal{B}_{22} = \{0\}$ , since

$$\begin{bmatrix} \mathbf{R}\mathcal{T} & -I \\ -I & 0 \end{bmatrix} \begin{bmatrix} z_4 \\ z_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

implies

$$\left. \begin{array}{l} \mathbf{R}\mathcal{T}z_4 - z_5 = 0 \\ -z_4 = 0 \end{array} \right\} \Rightarrow \begin{bmatrix} z_4 \\ z_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Thus, by Lemma 3.1 the semi-explicit abstract DAE defined by (5.15) for this problem is an index-1 DAE.

We investigate the image of  $\mathcal{B}_{22}$ :

$$\begin{bmatrix} \mathbf{R}\mathcal{T} & -I \\ -I & 0 \end{bmatrix} \begin{bmatrix} z_4 \\ z_5 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

implies

$$\begin{aligned} y_1 &= \mathbf{R}\mathcal{T}z_4 - z_5 \\ y_2 &= -z_4. \end{aligned}$$

Thus, we have  $y_2 \in \mathcal{H}_\sigma$ . Solving for  $z_4$  and  $z_5$  yields

$$\begin{bmatrix} z_4 \\ z_5 \end{bmatrix} = \begin{bmatrix} -y_2 \\ -y_1 - \mathbf{R}\mathcal{T}y_2 \end{bmatrix}.$$

It follows that,  $\text{im } \mathcal{B}_{22} = \mathbb{R}^4 \times \mathcal{H}_\sigma$  and  $\text{im } \mathcal{B}_{22}$  is dense in  $\mathcal{Y}_2$ . We have then  $\mathcal{D}(\mathcal{B}_{22})$  is dense in  $\{L^2(0, L) \times L^2(0, L)\} \times \{L^2(0, L) \times L^2(0, L)\} \times \mathbb{R}^3$  and  $\mathcal{B}_{22}$  is densely solvable. Since  $\mathcal{B}_{22}$  is one-to-one, we can define its inverse  $\mathcal{B}_{22}^{-1} : \mathbb{R}^4 \times \mathcal{H}_\sigma \rightarrow \mathcal{H}_\sigma \times \mathbb{R}^4$  as

$$\mathcal{B}_{22}^{-1} = \begin{bmatrix} 0 & -I \\ -I & \mathbf{R}\mathcal{T} \end{bmatrix}.$$

We next investigate the image of  $\mathcal{B}_{21}$  where

$$\mathcal{B}_{21} : \mathcal{D}(\mathcal{B}_{21}) = \{(z_1, z_2, z_3)^T \in \mathcal{H}_1 \times \mathcal{H}_2 \times \mathbb{R}^3 : \mathcal{T} z_2(t) = \mathbf{L}^T z_3(t)\} \rightarrow \mathcal{Y}_2:$$

$$\begin{bmatrix} 0 & 0 & 0 \\ \mathbf{E}D^2 & \mathbf{G}D^2 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

implies

$$\begin{aligned} y_1 &= 0, \\ y_2 &= \mathbf{E}D^2 z_1 + \mathbf{G}D^2 z_2. \end{aligned}$$

Hence,  $y_2 \in \mathcal{H}_\sigma$  and  $\text{im } \mathcal{B}_{21} = \{0\} \times \mathcal{H}_\sigma$  where  $0 \in \mathbb{R}^4$ .

Finally, we investigate the image of  $\mathcal{B}_{12}$  where

$$\mathcal{B}_{12} : \mathcal{H}_\sigma \times \mathbb{R}^4 \rightarrow \mathcal{Y}_1 = \{L^2(0, L) \times L^2(0, L)\} \times \{L^2(0, L) \times L^2(0, L)\} \times \mathbb{R}^3:$$

$$\begin{bmatrix} 0 & 0 \\ \mathbf{A}D^2 & 0 \\ 0 & -\mathbf{M}^{-1}\mathbf{L} \end{bmatrix} \begin{bmatrix} z_4 \\ z_5 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

implies

$$\begin{aligned} y_1 &= 0, \\ y_2 &= \mathbf{A}D^2 z_4, \\ y_3 &= -\mathbf{M}^{-1}\mathbf{L}z_5. \end{aligned}$$

We have,  $y_2 \in \{L^2(0, L) \times L^2(0, L)\}$ . Since  $\mathbf{L} \in \mathbb{R}^{3 \times 4}$  with full row rank and  $\mathbf{M}$  is full rank,  $\mathbf{M}^{-1}\mathbf{L}$  has full row rank. Therefore, we can solve for  $z_5$  above as  $z_5 = -(\mathbf{L}^T \mathbf{M}^{-1} \mathbf{L})^{-1} \mathbf{L}^T y_3$ . Thus, given any  $y_3 \in \mathbb{R}^3$  we can find a  $z_5 \in \mathbb{R}^4$ . Likewise for any  $y_2 \in \{L^2(0, L) \times L^2(0, L)\}$  we can find a  $z_4 \in \mathcal{H}_\sigma = \{H^2(0, L) \times H^2(0, L)\}$ . Hence,  $\text{im } \mathcal{B}_{12} = \{0\} \times \{L^2(0, L) \times L^2(0, L)\} \times \mathbb{R}^3$ .

From (3.3), we have

$$G_1 = \begin{bmatrix} I & \mathcal{B}_{12} \\ 0 & \mathcal{B}_{22} \end{bmatrix}.$$

We define  $\mathcal{D}(G_1) = \{L^2(0, L) \times L^2(0, L)\} \times \{L^2(0, L) \times L^2(0, L)\} \times \mathbb{R}^3 \times \mathcal{H}_\sigma \times \mathbb{R}^4$  which is dense in  $\mathcal{H}$ . Then,  $G_1 : \mathcal{D}(G_1) \subset \mathcal{H} \rightarrow \mathcal{Y}_1 \times \mathcal{Y}_2$ . We see that  $G_1$  is one-to-one as

$$\begin{bmatrix} I & \mathcal{B}_{12} \\ 0 & \mathcal{B}_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

implies

$$\left. \begin{aligned} x_1 + \mathcal{B}_{12}x_2 &= 0 \\ \mathcal{B}_{22}x_2 &= 0 \end{aligned} \right\} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

since  $\mathcal{B}_{22}$  is one-to-one. We next investigate  $\text{im } G_1$ :

$$\begin{bmatrix} I & \mathcal{B}_{12} \\ 0 & \mathcal{B}_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix},$$

where  $x_1 \in \{L^2(0, L) \times L^2(0, L)\} \times \{L^2(0, L) \times L^2(0, L)\} \times \mathbb{R}^3$  and  $x_2 \in \mathcal{D}(\mathcal{B}_{12}) = \mathcal{D}(\mathcal{B}_{22}) = \mathcal{H}_\sigma \times \mathbb{R}^4$ . This implies

$$\begin{aligned} x_1 + \mathcal{B}_{12}x_2 &= y_1, \\ \mathcal{B}_{22}x_2 &= y_2. \end{aligned}$$

However, we have  $\text{im } \mathcal{B}_{12} = \{0\} \times \{L^2(0, L) \times L^2(0, L)\} \times \mathbb{R}^3$ , and  $\text{im } \mathcal{B}_{22} = \mathbb{R}^4 \times \mathcal{H}_\sigma$ . Thus,  $y_2 \in \mathbb{R}^4 \times \mathcal{H}_\sigma$  and  $y_1 \in \{L^2(0, L) \times L^2(0, L)\} \times \{L^2(0, L) \times L^2(0, L)\} \times \mathbb{R}^3$ . Therefore,  $\text{im } G_1$  is dense in  $\mathcal{Y}_1 \times \mathcal{Y}_2$ . Hence,  $G_1$  is densely defined, densely solvable, and one-to-one. This is exactly the minimum requirement for a DAE to be index-1 by Definition 3.2.

## 5.2.2 Existence and Uniqueness

Now that we are done setting up and analyzing the problem, we show that there exists a unique solution to this problem using the general theory developed for abstract index-1 DAEs in Chapter 3.

**Theorem 5.1.** *Given consistent index-1 initial condition  $w(0) = [f_1 \ f_2]^T$  and the right-hand side function  $q(t)$  Lipschitz continuous, there exists a unique solution  $z \in C([0, T]; \mathcal{H})$  for this two beam and joint problem.*

*Proof.* We proceed to the AODE using the direct method. From (5.15) we have two equations which can be used to write the AODE in terms of  $u(t) = x_1(t) \in \mathcal{D}(\mathcal{B}_{11}) = \mathcal{D}(\mathcal{B}_{21}) = \{x_1 = (z_1, z_2, z_3)^T \in \mathcal{H}_1 \times \mathcal{H}_2 \times \mathbb{R}^3 : \mathcal{T}z_2(t) = \mathbf{L}^T z_3(t)\}$ :

$$\begin{aligned} x_1'(t) + \mathcal{B}_{11}x_1(t) + \mathcal{B}_{12}x_2(t) &= q_1(t), \\ \mathcal{B}_{21}x_1(t) + \mathcal{B}_{22}x_2(t) &= 0 \end{aligned}$$

implies

$$u'(t) + (\mathcal{B}_{11} - \mathcal{B}_{12}\mathcal{B}_{22}^{-1}\mathcal{B}_{21})u(t) = q_1(t),$$

where we used the constraint equation along with the fact that  $\mathcal{B}_{22}$  is one-to-one to obtain  $x_2(t) = -\mathcal{B}_{22}^{-1}\mathcal{B}_{21}x_1(t)$ . This expression makes sense as  $\mathcal{D}(\mathcal{B}_{22}^{-1}) \supset \text{im } \mathcal{B}_{21}$  and  $\mathcal{D}(\mathcal{B}_{12}) \supseteq \text{im } \mathcal{B}_{22}^{-1}$ .

We let  $\mathcal{A} = -(\mathcal{B}_{11} - \mathcal{B}_{12}\mathcal{B}_{22}^{-1}\mathcal{B}_{21})$  and rewrite our AODE as:

$$u'(t) = \mathcal{A}u(t) + q_1(t). \tag{5.16}$$

The definitions for the block operator matrices and calculating  $-(\mathcal{B}_{11} - \mathcal{B}_{12}\mathcal{B}_{22}^{-1}\mathcal{B}_{21})$  yields

$$\mathcal{A} = \begin{bmatrix} 0 & I & 0 \\ -\mathbf{A}\mathbf{E}D^4 & -\mathbf{A}\mathbf{G}D^4 & 0 \\ \mathbf{M}^{-1}\mathbf{L}\mathbf{R}\mathbf{T}\mathbf{E}D^2 & \mathbf{M}^{-1}\mathbf{L}\mathbf{R}\mathbf{T}\mathbf{G}D^2 & -\mathbf{M}^{-1}\mathbf{W} \end{bmatrix}. \quad (5.17)$$

We define  $\mathcal{H}_{ODE} = \mathcal{H}_1 \times \{L^2(0, L) \times L^2(0, L)\} \times \mathbb{R}^3$  and note that  $\mathcal{D}(\mathcal{A}) = \{(z_1, z_2, z_3)^T \in \mathcal{H}_1 \times \mathcal{H}_2 \times \mathbb{R}^3 : \mathbf{T}z_2 = \mathbf{L}^T z_3\}$ . Furthermore,  $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subset \mathcal{H}_{ODE} \rightarrow \mathcal{H}_{ODE}$  and  $\text{im } \mathcal{A} = \mathcal{H}_{ODE}$ . We have  $\text{im } \mathcal{B}_{11} = \mathcal{H}_1 \times \{0\} \times \text{im } (\mathbf{M}^{-1}\mathbf{W})$  and  $\text{im } \mathcal{B}_{12} = \{0\} \times \{L^2(0, L) \times L^2(0, L)\} \times \mathbb{R}^3$ . Then,  $\text{im } \mathcal{A} = \mathcal{B}_{11} - \mathcal{B}_{12}\mathcal{B}_{22}^{-1}\mathcal{B}_{21} = \text{im } \mathcal{B}_{11} \cup \text{im } \mathcal{B}_{12} = \mathcal{H}_{ODE}$ . We will formally verify that  $\mathcal{A}$  maps onto  $\mathcal{H}_{ODE}$  below.

We first define an energy norm on  $\mathcal{H}_{ODE}$  as

$$\|z\|_{\mathcal{H}_{ODE}}^2 = (\mathbf{E}D^2 z_1, D^2 z_1)_{L^2(0, L)^2} + (\mathbf{A}^{-1} z_2, z_2)_{L^2(0, L)^2} + z_3^T \mathbf{M} z_3,$$

or if we use our original variables

$$\|z\|_{\mathcal{H}_{ODE}}^2 = (\mathbf{E}D^2 w, D^2 w)_{L^2(0, L)^2} + (\mathbf{A}^{-1} w', w')_{L^2(0, L)^2} + (y')^T \mathbf{M} y',$$

where, in an abuse of notation, we use  $L^2(0, L)^2 = L^2(0, L) \times L^2(0, L)$ . Also,  $\mathbf{A}^{-1}$  is just the positive definite symmetric matrix

$$\mathbf{A}^{-1} = \begin{bmatrix} \rho A & 0 \\ 0 & \rho A \end{bmatrix}.$$

At this point we want to show that  $\mathcal{A}$  is the infinitesimal generator of a strongly continuous semigroup of contractions. To do so we use a Corollary to the Lumer-Phillips Theorem found in Liu & Zheng [30]. This Corollary states that if  $\mathcal{A}$  is a linear operator with dense domain and is dissipative with zero in its resolvent set,  $\rho(\mathcal{A})$ , then  $\mathcal{A}$  is the infinitesimal generator of a strongly continuous semigroup of contractions.

$\mathcal{A}$  is linear with dense domain in  $\mathcal{H}_{ODE}$ . We now show that  $\mathcal{A}$  is dissipative, i.e.,  $\text{Re}(\mathcal{A}z, z) \leq 0$ .

Before we calculate  $(\mathcal{A}z, z)_{\mathcal{H}_{ODE}}$  we will first need  $\mathcal{A}z$ . We will need to do this calculation with the original variables  $z = (w, w', y')^T$ :

$$\begin{aligned} \mathcal{A}z &= \begin{bmatrix} 0 & I & 0 \\ -\mathbf{A}\mathbf{E}D^4 & -\mathbf{A}\mathbf{G}D^4 & 0 \\ \mathbf{M}^{-1}\mathbf{L}\mathbf{R}\mathbf{T}\mathbf{E}D^2 & \mathbf{M}^{-1}\mathbf{L}\mathbf{R}\mathbf{T}\mathbf{G}D^2 & -\mathbf{M}^{-1}\mathbf{W} \end{bmatrix} \begin{bmatrix} w \\ w' \\ y' \end{bmatrix} \\ &= \begin{bmatrix} w' \\ -\mathbf{A}\mathbf{E}D^4 w - \mathbf{A}\mathbf{G}D^4 w' \\ \mathbf{M}^{-1}\mathbf{L}\mathbf{R}\mathbf{T}\mathbf{E}D^2 w + \mathbf{M}^{-1}\mathbf{L}\mathbf{R}\mathbf{T}\mathbf{G}D^2 w' - \mathbf{M}^{-1}\mathbf{W} y' \end{bmatrix}. \end{aligned}$$

We are now ready to calculate this inner product:

$$(\mathcal{A}z, z)_{\mathcal{H}_{ODE}} = (\mathbf{E}D^2w', D^2w)_{L^2(0,L)^2} \quad (5.18)$$

$$+ (\mathbf{A}^{-1}(-\mathbf{A}\mathbf{E}D^4w - \mathbf{A}\mathbf{G}D^4w'), w')_{L^2(0,L)^2} \quad (5.19)$$

$$+ \underbrace{(y')^T \mathbf{M}(\mathbf{M}^{-1}\mathbf{L}\mathbf{R}\mathbf{T}\mathbf{E}D^2w + \mathbf{M}^{-1}\mathbf{L}\mathbf{R}\mathbf{T}\mathbf{G}D^2w' - \mathbf{M}^{-1}\mathbf{W}y')}_{= IP_1(5.18) + IP_2(5.19) + IP_3(5.20)}. \quad (5.20)$$

We concentrate on the first two terms  $IP_1$  and  $IP_2$ :

$$\begin{aligned} & (\mathbf{E}D^2w', D^2w)_{L^2(0,L)^2} + (\mathbf{A}^{-1}(-\mathbf{A}\mathbf{E}D^4w - \mathbf{A}\mathbf{G}D^4w'), w')_{L^2(0,L)^2} \\ &= (\mathbf{E}D^2w', D^2w)_{L^2(0,L)^2} - (\mathbf{E}D^4w, w')_{L^2(0,L)^2} - (\mathbf{G}D^4w', w'(t))_{L^2(0,L)^2}. \end{aligned}$$

Next using the definition of the inner product in the product space  $L^2(0, L) \times L^2(0, L)$  and integrating by parts, we have

$$\begin{aligned} (\mathbf{E}D^4w, w')_{L^2(0,L)^2} &= \left( \begin{bmatrix} EI \frac{\partial^4 w_1}{\partial s^4} \\ EI \frac{\partial^4 w_2}{\partial s^4} \end{bmatrix}, \begin{bmatrix} w'_1 \\ w'_2 \end{bmatrix} \right)_{L^2(0,L)^2} \\ &= EI \int_0^L \frac{\partial^4 w_1}{\partial s^4} w'_1 dx + EI \int_0^L \frac{\partial^4 w_2}{\partial s^4} w'_2 dx \\ &= EI \frac{\partial^3 w_1}{\partial s^3} w'_1 \Big|_0^L - EI \int_0^L \frac{\partial^3 w_1}{\partial s^3} \frac{\partial w'_1}{\partial s} dx \\ &\quad + EI \frac{\partial^3 w_2}{\partial s^3} w'_2 \Big|_0^L - EI \int_0^L \frac{\partial^3 w_2}{\partial s^3} \frac{\partial w'_2}{\partial s} dx \\ &= EI \frac{\partial^3 w_1}{\partial s^3}(L) w'_1(L) - EI \frac{\partial^2 w_1}{\partial s^2} \frac{\partial w'_1}{\partial s} \Big|_0^L + EI \int_0^L \frac{\partial^2 w_1}{\partial s^2} \frac{\partial^2 w'_1}{\partial s^2} dx \\ &\quad + EI \frac{\partial^3 w_2}{\partial s^3}(L) w'_2(L) - EI \frac{\partial^2 w_2}{\partial s^2} \frac{\partial w'_2}{\partial s} \Big|_0^L + EI \int_0^L \frac{\partial^2 w_2}{\partial s^2} \frac{\partial^2 w'_2}{\partial s^2} dx \\ &= EI \frac{\partial^3 w_1}{\partial s^3}(L) w'_1(L) - EI \frac{\partial^2 w_1}{\partial s^2}(L) \frac{\partial w'_1}{\partial s}(L) + EI \frac{\partial^3 w_2}{\partial s^3}(L) w'_2(L) \\ &\quad - EI \frac{\partial^2 w_2}{\partial s^2}(L) \frac{\partial w'_2}{\partial s}(L) + (\mathbf{E}D^2w', D^2w)_{L^2(0,L)^2}. \quad (5.21) \end{aligned}$$

In evaluating the functions at the end  $s = 0$  we have used the fact that  $w'_i \in H_\ell^2(0, L)$ .

Similarly we have

$$\begin{aligned}
(\mathbf{G}D^4w', w')_{L^2(0,L)^2} &= \left( \begin{bmatrix} \gamma \frac{\partial^4 w'_1}{\partial s^4} \\ \gamma \frac{\partial^4 w'_2}{\partial s^4} \end{bmatrix}, \begin{bmatrix} w'_1 \\ w'_2 \end{bmatrix} \right)_{L^2(0,L)^2} \\
&= \gamma \int_0^L \frac{\partial^4 w'_1}{\partial s^4} w'_1 dx + \gamma \int_0^L \frac{\partial^4 w'_2}{\partial s^4} w'_2 dx \\
&= \gamma \frac{\partial^3 w'_1}{\partial s^3} w'_1 \Big|_0^L - \gamma \int_0^L \frac{\partial^3 w'_1}{\partial s^3} \frac{\partial w'_1}{\partial s} dx \\
&\quad + \gamma \frac{\partial^3 w'_2}{\partial s^3} w'_2 \Big|_0^L - \gamma \int_0^L \frac{\partial^3 w'_2}{\partial s^3} \frac{\partial w'_2}{\partial s} dx \\
&= \gamma \frac{\partial^3 w'_1}{\partial s^3}(L) w'_1(L) - \gamma \frac{\partial^2 w'_1}{\partial s^2} \frac{\partial w'_1}{\partial s} \Big|_0^L + \gamma \int_0^L \frac{\partial^2 w'_1}{\partial s^2} \frac{\partial^2 w'_1}{\partial s^2} dx \\
&\quad + \gamma \frac{\partial^3 w'_2}{\partial s^3}(L) w'_2(L) - \gamma \frac{\partial^2 w'_2}{\partial s^2} \frac{\partial w'_2}{\partial s} \Big|_0^L + \gamma \int_0^L \frac{\partial^2 w'_2}{\partial s^2} \frac{\partial^2 w'_2}{\partial s^2} dx \\
&= \gamma \frac{\partial^3 w'_1}{\partial s^3}(L) w'_1(L) - \gamma \frac{\partial^2 w'_1}{\partial s^2}(L) \frac{\partial w'_1}{\partial s}(L) + \gamma \frac{\partial^3 w'_2}{\partial s^3}(L) w'_2(L) \\
&\quad - \gamma \frac{\partial^2 w'_2}{\partial s^2}(L) \frac{\partial w'_2}{\partial s}(L) + (\mathbf{G}D^2w', D^2w')_{L^2(0,L)^2} \\
&= \gamma \frac{\partial^3 w'_1}{\partial s^3}(L) w'_1(L) - \gamma \frac{\partial^2 w'_1}{\partial s^2}(L) \frac{\partial w'_1}{\partial s}(L) + \gamma \frac{\partial^3 w'_2}{\partial s^3}(L) w'_2(L) \\
&\quad - \gamma \frac{\partial^2 w'_2}{\partial s^2}(L) \frac{\partial w'_2}{\partial s}(L) + (\mathbf{G}^{\frac{1}{2}}D^2w', \mathbf{G}^{\frac{1}{2}}D^2w')_{L^2(0,L)^2}, \quad (5.22)
\end{aligned}$$

where we use that  $\mathbf{G}$  is positive definite symmetric

$$\mathbf{G} = \begin{bmatrix} \gamma & 0 \\ 0 & \gamma \end{bmatrix} = \begin{bmatrix} \gamma^{\frac{1}{2}} & 0 \\ 0 & \gamma^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \gamma^{\frac{1}{2}} & 0 \\ 0 & \gamma^{\frac{1}{2}} \end{bmatrix} = \mathbf{G}^{\frac{1}{2}} \mathbf{G}^{\frac{1}{2}}.$$

Now using (5.21) and (5.22), we have for  $IP_1 + IP_2$ :

$$\begin{aligned}
IP_1 + IP_2 &= (\mathbf{E}D^2w', D^2w)_{L^2(0,L)^2} - (\mathbf{E}D^4w, w')_{L^2(0,L)^2} - (\mathbf{G}D^4w', w'(t))_{L^2(0,L)^2} \\
&= (\mathbf{E}D^2w', D^2w)_{L^2(0,L)^2} - EI \frac{\partial^3 w_1}{\partial s^3}(L)w_1'(L) + EI \frac{\partial^2 w_1}{\partial s^2}(L) \frac{\partial w_1'}{\partial s}(L) \\
&\quad - EI \frac{\partial^3 w_2}{\partial s^3}(L)w_2'(L) + EI \frac{\partial^2 w_2}{\partial s^2}(L) \frac{\partial w_2'}{\partial s}(L) - (\mathbf{E}D^2w', D^2w)_{L^2(0,L)^2} \\
&\quad - \gamma \frac{\partial^3 w_1'}{\partial s^3}(L)w_1'(L) + \gamma \frac{\partial^2 w_1'}{\partial s^2}(L) \frac{\partial w_1'}{\partial s}(L) - \gamma \frac{\partial^3 w_2'}{\partial s^3}(L)w_2'(L) \\
&\quad + \gamma \frac{\partial^2 w_2'}{\partial s^2}(L) \frac{\partial w_2'}{\partial s}(L) - (\mathbf{G}^{\frac{1}{2}}D^2w', \mathbf{G}^{\frac{1}{2}}D^2w')_{L^2(0,L)^2} \\
&= \left[ EI \frac{\partial^2 w_1}{\partial s^2}(L) + \gamma \frac{\partial^2 w_1'}{\partial s^2}(L) \right] \frac{\partial w_1'}{\partial s}(L) + \left[ EI \frac{\partial^2 w_2}{\partial s^2}(L) + \gamma \frac{\partial^2 w_2'}{\partial s^2}(L) \right] \frac{\partial w_2'}{\partial s}(L) \\
&\quad - \left[ EI \frac{\partial^3 w_1}{\partial s^3}(L) + \gamma \frac{\partial^3 w_1'}{\partial s^3}(L) \right] w_1'(L) - \left[ EI \frac{\partial^3 w_2}{\partial s^3}(L) + \gamma \frac{\partial^3 w_2'}{\partial s^3}(L) \right] w_2'(L) \\
&\quad - \|\mathbf{G}^{\frac{1}{2}}D^2w'\|_{L^2(0,L)^2} \\
&= M_1 \frac{\partial w_1'}{\partial s}(L) + M_2 \frac{\partial w_2'}{\partial s}(L) - N_1 w_1'(L) - N_2 w_2'(L) - \|\mathbf{G}^{\frac{1}{2}}D^2w'\|_{L^2(0,L)^2}.
\end{aligned} \tag{5.23}$$

We next investigate the term  $IP_3$ :

$$\begin{aligned}
IP_3 &= (y')^T \mathbf{M}(\mathbf{M}^{-1} \mathbf{LRTED}^2w + \mathbf{M}^{-1} \mathbf{LRTGD}^2w' - \mathbf{M}^{-1} \mathbf{W}y') \\
&= (y')^T \mathbf{LRTED}^2w + (y')^T \mathbf{LRTGD}^2w' - (y')^T \mathbf{W}y'.
\end{aligned}$$

We look at each of these three terms separately. The first term yields

$$\begin{aligned}
(y')^T \mathbf{LRTED}^2w &= [\dot{x} \quad \dot{\theta}_1 \quad \dot{\theta}_2] \begin{bmatrix} 0 & -1 & 0 & 1 \\ 1 & \ell & 0 & 0 \\ 0 & 0 & 1 & \ell \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} -EI \frac{\partial^3 w_1}{\partial s^3}(L) \\ EI \frac{\partial^2 w_1}{\partial s^2}(L) \\ -EI \frac{\partial^3 w_2}{\partial s^3}(L) \\ EI \frac{\partial^2 w_2}{\partial s^2}(L) \end{bmatrix} \\
&= -EI \frac{\partial^3 w_1}{\partial s^3}(L)(\dot{x} - \ell \dot{\theta}_1) + EI \frac{\partial^3 w_2}{\partial s^3}(L)(\dot{x} + \ell \dot{\theta}_2) \\
&\quad + EI \frac{\partial^2 w_1}{\partial s^2}(L)\dot{\theta}_1 + EI \frac{\partial^2 w_2}{\partial s^2}(L)\dot{\theta}_2.
\end{aligned} \tag{5.24}$$



Similarly, the second term yields

$$\begin{aligned}
(y')^T \mathbf{LRTGD}^2 w' &= [\dot{x} \quad \dot{\theta}_1 \quad \dot{\theta}_2] \begin{bmatrix} 0 & -1 & 0 & 1 \\ 1 & \ell & 0 & 0 \\ 0 & 0 & 1 & \ell \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} -\gamma \frac{\partial^3 w'_1}{\partial s^3}(L) \\ \gamma \frac{\partial^2 w'_1}{\partial s^2}(L) \\ -\gamma \frac{\partial^3 w'_2}{\partial s^3}(L) \\ \gamma \frac{\partial^2 w'_2}{\partial s^2}(L) \end{bmatrix} \\
&= -\gamma \frac{\partial^3 w'_1}{\partial s^3}(L)(\dot{x} - \ell \dot{\theta}_1) + \gamma \frac{\partial^3 w'_2}{\partial s^3}(L)(\dot{x} + \ell \dot{\theta}_2) \\
&\quad + \gamma \frac{\partial^2 w'_1}{\partial s^2}(L)\dot{\theta}_1 + \gamma \frac{\partial^2 w'_2}{\partial s^2}(L)\dot{\theta}_2. \tag{5.25}
\end{aligned}$$

Now we also must satisfy the domain constraint  $\mathcal{T}w' = \mathbf{L}^T y'$ , or

$$\begin{bmatrix} -\frac{\partial w'_1}{\partial s}(L) \\ w'_1(L) \\ -\frac{\partial w'_2}{\partial s}(L) \\ w'_2(L) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & \ell & 0 \\ 0 & 0 & 1 \\ 1 & 0 & \ell \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix} = \begin{bmatrix} \dot{\theta}_1 \\ -(\dot{x} - \ell \dot{\theta}_1) \\ \dot{\theta}_2 \\ (\dot{x} + \ell \dot{\theta}_2) \end{bmatrix}. \tag{5.26}$$

Substituting (5.26) into (5.24) yields

$$\begin{aligned}
(y')^T \mathbf{LRTED}^2 w &= EI \frac{\partial^3 w_1}{\partial s^3}(L)w'_1(L) + EI \frac{\partial^3 w_2}{\partial s^3}(L)w'_2(L) \\
&\quad - EI \frac{\partial^2 w_1}{\partial s^2}(L) \frac{\partial w'_1}{\partial s}(L) - EI \frac{\partial^2 w_2}{\partial s^2}(L) \frac{\partial w'_2}{\partial s}(L). \tag{5.27}
\end{aligned}$$

Similarly, substituting (5.26) into (5.25) yields

$$\begin{aligned}
(y')^T \mathbf{LRTGD}^2 w' &= \gamma \frac{\partial^3 w'_1}{\partial s^3}(L)w'_1(L) + \gamma \frac{\partial^3 w'_2}{\partial s^3}(L)w'_2(L) \\
&\quad - \gamma \frac{\partial^2 w'_1}{\partial s^2}(L) \frac{\partial w'_1}{\partial s}(L) - \gamma \frac{\partial^2 w'_2}{\partial s^2}(L) \frac{\partial w'_2}{\partial s}(L). \tag{5.28}
\end{aligned}$$

Combining (5.27) and (5.28) we have

$$\begin{aligned}
& (y')^T \mathbf{LRTED}^2 w + (y')^T \mathbf{LRTGD}^2 w' \\
&= - \left[ EI \frac{\partial^2 w_1}{\partial s^2}(L) + \gamma \frac{\partial^2 w_1'}{\partial s^2}(L) \right] \frac{\partial w_1'}{\partial s}(L) - \left[ EI \frac{\partial^2 w_2}{\partial s^2}(L) + \gamma \frac{\partial^2 w_2'}{\partial s^2}(L) \right] \frac{\partial w_2'}{\partial s}(L) \\
&\quad + \left[ EI \frac{\partial^3 w_1}{\partial s^3}(L) + \gamma \frac{\partial^3 w_1'}{\partial s^3}(L) \right] w_1'(L) + \left[ EI \frac{\partial^3 w_2}{\partial s^3}(L) + \gamma \frac{\partial^3 w_2'}{\partial s^3}(L) \right] w_2'(L) \\
&= -M_1 \frac{\partial w_1'}{\partial s}(L) - M_2 \frac{\partial w_2'}{\partial s}(L) + N_1 w_1'(L) + N_2 w_2'(L). \tag{5.29}
\end{aligned}$$

Finally, for the third term we get

$$\begin{aligned}
-(y')^T \mathbf{W} y' &= - \begin{bmatrix} \dot{x} & \dot{\theta}_1 & \dot{\theta}_2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & b & -b \\ 0 & -b & b \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix} \\
&= -b(\dot{\theta}_1 - \dot{\theta}_2)^2, \tag{5.30}
\end{aligned}$$

where  $b > 0$ . Thus, putting it all together with equations (5.20), (5.23), (5.29) and (5.30) we end up with

$$\begin{aligned}
(\mathcal{A}z, z)_{\mathcal{H}_{ODE}} &= M_1 \frac{\partial w_1'}{\partial s}(L) + M_2 \frac{\partial w_2'}{\partial s}(L) - N_1 w_1'(L) - N_2 w_2'(L) - \|\mathbf{G}^{\frac{1}{2}} D^2 w'\|_{L^2(0,L)^2} \\
&\quad - M_1 \frac{\partial w_1'}{\partial s}(L) - M_2 \frac{\partial w_2'}{\partial s}(L) + N_1 w_1'(L) + N_2 w_2'(L) - b(\dot{\theta}_1 - \dot{\theta}_2)^2 \\
&= -\|\mathbf{G}^{\frac{1}{2}} D^2 w'\|_{L^2(0,L)^2} - b(\dot{\theta}_1 - \dot{\theta}_2)^2 \leq 0. \tag{5.31}
\end{aligned}$$

Therefore, we can conclude that  $\mathcal{A}$  is dissipative.

We next work to show that zero is in the resolvent set of  $\mathcal{A}$ . We do this by showing that  $\mathcal{A}$  is one-to-one and onto. Then,  $\mathcal{A}^{-1}$  exists and is bounded.

We set  $\mathcal{A}z = \tilde{z}$ :

$$\begin{aligned}
\mathcal{A}z &= \begin{bmatrix} 0 & I & 0 \\ -\mathbf{AED}^4 & -\mathbf{AGD}^4 & 0 \\ \mathbf{M}^{-1}\mathbf{LRTED}^2 & \mathbf{M}^{-1}\mathbf{LRTGD}^2 & -\mathbf{M}^{-1}\mathbf{W} \end{bmatrix} \begin{bmatrix} w \\ w' \\ y' \end{bmatrix} \\
&= \begin{bmatrix} w' \\ -\mathbf{AED}^4 w - \mathbf{AGD}^4 w' \\ \mathbf{M}^{-1}\mathbf{LRTED}^2 w + \mathbf{M}^{-1}\mathbf{LRTGD}^2 w' - \mathbf{M}^{-1}\mathbf{W} y' \end{bmatrix} = \begin{bmatrix} \tilde{z}_1 \\ \tilde{z}_2 \\ \tilde{z}_3 \end{bmatrix}.
\end{aligned}$$

This gives three equations:

$$w' = \tilde{z}_1 \in \mathcal{H}_1, \tag{5.32}$$

$$-\mathbf{AED}^4 w - \mathbf{AGD}^4 w' = \tilde{z}_2 \in \{L^2(0, L) \times L^2(0, L)\}, \tag{5.33}$$

$$\mathbf{M}^{-1}\mathbf{LRTED}^2 w + \mathbf{M}^{-1}\mathbf{LRTGD}^2 w' - \mathbf{M}^{-1}\mathbf{W} y' = \tilde{z}_3 \in \mathbb{R}^3. \tag{5.34}$$

Equation (5.32) yields a unique  $w'$  for given  $\tilde{z}_1$ . We then substitute for  $w'$  in (5.33) and solve for  $w$  to obtain an ODE with respect to  $s$ ,

$$\begin{aligned} -\mathbf{A}\mathbf{E}D^4w &= \tilde{z}_2 + \mathbf{A}\mathbf{G}D^4\tilde{z}_1 \text{ or,} \\ D^4w &= -\mathbf{E}^{-1}\mathbf{A}^{-1}\tilde{z}_2 - \mathbf{E}^{-1}\mathbf{G}D^4\tilde{z}_1. \end{aligned} \quad (5.35)$$

From the domain constraint for  $\mathcal{A}$  we have  $\mathcal{T}w' = \mathbf{L}^T y'$ , or  $y' = (\mathbf{L}\mathbf{L}^T)^{-1}\mathbf{L}\mathcal{T}w'$ , or  $y' = (\mathbf{L}\mathbf{L}^T)^{-1}\mathbf{L}\mathcal{T}\tilde{z}_1$  where we used (5.32). Substituting for  $y'$  and  $w'$  into (5.34) yields

$$\begin{aligned} \mathbf{M}^{-1}\mathbf{L}\mathbf{R}\mathcal{T}\mathbf{E}D^2w &= \tilde{z}_3 - \mathbf{M}^{-1}\mathbf{L}\mathbf{R}\mathcal{T}\mathbf{G}D^2\tilde{z}_1 + \mathbf{M}^{-1}\mathbf{W}(\mathbf{L}\mathbf{L}^T)^{-1}\mathbf{L}\mathcal{T}\tilde{z}_1 \\ \mathbf{L}\mathbf{R}\mathcal{T}\mathbf{E}D^2w &= \mathbf{M}\tilde{z}_3 - \mathbf{L}\mathbf{R}\mathcal{T}\mathbf{G}D^2\tilde{z}_1 + \mathbf{W}(\mathbf{L}\mathbf{L}^T)^{-1}\mathbf{L}\mathcal{T}\tilde{z}_1 \\ \mathbf{R}\mathcal{T}\mathbf{E}D^2w &= (\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T\mathbf{M}\tilde{z}_3 - \mathbf{R}\mathcal{T}\mathbf{G}D^2\tilde{z}_1 + (\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T\mathbf{W}(\mathbf{L}\mathbf{L}^T)^{-1}\mathbf{L}\mathcal{T}\tilde{z}_1 \\ \mathcal{T}\mathbf{E}D^2w &= \mathbf{R}^{-1}(\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T\mathbf{M}\tilde{z}_3 - \mathcal{T}\mathbf{G}D^2\tilde{z}_1 + \mathbf{R}^{-1}(\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T\mathbf{W}(\mathbf{L}\mathbf{L}^T)^{-1}\mathbf{L}\mathcal{T}\tilde{z}_1, \end{aligned}$$

where we used the Moore-Penrose pseudoinverse of  $\mathbf{L}$ , i.e.,  $(\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T$ , to solve for  $\mathcal{T}\mathbf{E}D^2w$ .

We write the right-hand side as known functions and calculate the left-hand side:

$$\begin{aligned} \begin{bmatrix} -EI \frac{d^3w_1}{ds^3}(L) \\ EI \frac{d^2w_1}{ds^2}(L) \\ -EI \frac{d^3w_2}{ds^3}(L) \\ EI \frac{d^2w_2}{ds^2}(L) \end{bmatrix} &= \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix}, \text{ or} \\ \begin{bmatrix} \frac{d^3w_1}{ds^3}(L) \\ \frac{d^2w_1}{ds^2}(L) \\ \frac{d^3w_2}{ds^3}(L) \\ \frac{d^2w_2}{ds^2}(L) \end{bmatrix} &= \begin{bmatrix} \tilde{f}_1 \\ \tilde{f}_2 \\ \tilde{f}_3 \\ \tilde{f}_4 \end{bmatrix}. \end{aligned} \quad (5.36)$$

Thus, equation (5.34) provides boundary conditions (5.36) for the fourth order ODE (5.35). Additionally, we also have the boundary conditions imposed by the domain of  $\mathcal{A}$ , i.e.,  $H_\ell^2(0, L)$  of  $w(0) = Dw(0) = 0$ . Thus, the lower order derivative terms provide boundary conditions at  $s = 0$  while the two higher order derivative terms provide

boundary conditions at  $s = L$ . By standard ODE theory, these particular boundary conditions result in existence of a unique solution for  $w$  (see for example Walter [40] Chapter VI, Section 26, Parts X and XI). Also by standard linear elliptic differential equation theory we have for a right-hand side function in  $\{L^2(0, L) \times L^2(0, L)\}$  that  $w \in \{H_\ell^2(0, L) \cap H^4(0, L)\} \times \{H_\ell^2(0, L) \cap H^4(0, L)\} = \mathcal{H}_1$ . We have established that given a right-hand side in  $\mathcal{H}_{ODE}$  there exists a unique solution of  $\mathcal{A}z = \tilde{z}$  for  $w \in \mathcal{H}_1$ ,  $w' \in \mathcal{H}_1 = \mathcal{H}_2$  and  $y' \in \mathbb{R}^3$  where  $y'$  satisfies  $\mathbf{L}^T y' = \mathcal{T}w'$ . As a result we have shown that  $\mathcal{A}^{-1}$  exists.

We now show that  $\mathcal{A}^{-1}$  is bounded. From expressions (5.32) and (5.35) and the expression for  $y'$  from the constraint equation, we have

$$\begin{aligned} \|w'\| &\leq c_1 \|\tilde{z}_1\| \\ \|w\| &\leq c_2 (\|\tilde{z}_2\| + \|\tilde{z}_1\|) \\ \|y'\| &\leq c_3 \|\tilde{z}_1\|. \end{aligned}$$

Thus,

$$\begin{aligned} \|z\|_{\mathcal{H}_{ODE}} &\leq C \|\tilde{z}_1\|_{\mathcal{H}_{ODE}}, \\ \|z\|_{\mathcal{H}_{ODE}} &\leq C \|\mathcal{A}z\|_{\mathcal{H}_{ODE}}, \\ \|\mathcal{A}^{-1}(\mathcal{A}z)\|_{\mathcal{H}_{ODE}} &\leq C \|\mathcal{A}z\|_{\mathcal{H}_{ODE}}. \end{aligned}$$

Therefore,  $\mathcal{A}^{-1}$  is bounded and the resolvent operator for  $\lambda = 0$ , i.e.,  $(0 - \mathcal{A})^{-1}$ , is also bounded. Thus, we have  $0 \in \rho(\mathcal{A})$ .

Hence, by the Corollary to Lumer-Phillips in Liu and Zheng [30] mentioned above,  $\mathcal{A}$  is the infinitesimal generator of a strongly continuous semigroup of contractions on  $\mathcal{H}_{ODE}$ .

Now by Lemma 3.7 with  $q(t)$  Lipschitz continuous, and hence also  $q_1(t)$ , we have a unique solution  $u(t) = x_1(t) \in \mathcal{D}(\mathcal{B}_{11}) = \mathcal{D}(\mathcal{B}_{21})$  to the inhomogeneous abstract Cauchy problem (5.16). Then using the expression  $x_2(t) = -\mathcal{B}_{22}^{-1} \mathcal{B}_{21} x_1(t)$  we obtain a unique solution for  $x_2(t) \in \mathcal{D}(\mathcal{B}_{12}) = \mathcal{D}(\mathcal{B}_{22})$ . We have therefore shown that there exists a unique solution to this two beam and rigid joint problem.  $\square$

# Chapter 6

## Discretized Differential-Algebraic Equations

Once we have shown that there exists a unique solution to the abstract DAE we then want to be able to develop numerical methods for approximating the solutions. One approach is to discretize the spatial variables of the abstract DAE using a finite element approach to form a finite dimensional DAE. We assume the boundary conditions and constraints that are imposed in the domain of the operators have been incorporated into the basis elements of the finite element space. In this chapter, we investigate some of the issues that arise after discretization such as solvability, projections, index reduction, etc.

### 6.1 Solvability of the Differential-Algebraic Equation

Consider the linear constant coefficient Differential-Algebraic Equation (DAE) that may be typical of those which arise from various structural dynamics applications:

$$\begin{cases} M\ddot{z}(t) + D\dot{z}(t) + Kz(t) - CF(t) = f(t), \\ C^T z(t) = 0, \end{cases} \quad (6.1)$$

where  $M \in \mathbb{R}^{n \times n}$  is a mass matrix with the property  $M > 0$  and symmetric,  $D \in \mathbb{R}^{n \times n}$  is a damping matrix,  $K \in \mathbb{R}^{n \times n}$  is a stiffness matrix with the usual property  $K \geq 0$  and symmetric,  $C \in \mathbb{R}^{n \times p}$  with full rank  $p$ ,  $p \leq n$ ,  $z(t), f(t) \in \mathbb{R}^n$  where  $f(t)$  is an optional forcing function, and  $F(t) \in \mathbb{R}^p$  are unknown functions related to the problem (For example these could be the internal forces and moments imposed on a joint by two vibrating beams like our application in Chapter 5). The second of the two equations above represents a constraint on the second-order ODE in the first equation. Note that

this system is a DAE because the unknown variable  $F(t)$  appears only in algebraic form and is not differentiated in the equation.

We now convert this second-order system into a first-order ODE. Let  $x_1(t) = z(t)$ ,  $x_2(t) = \dot{z}(t) = \dot{x}_1(t)$ , and  $x_3(t) = F(t)$ . We then get three equations:

$$\begin{aligned} \dot{x}_1(t) - x_2(t) &= 0, \\ M\dot{x}_2(t) + Dx_2(t) + Kx_1(t) - Cx_3(t) &= f(t), \\ C^T x_1(t) &= 0. \end{aligned}$$

We incorporate these three equations into matrix form,

$$\begin{bmatrix} I_n & 0_n & 0_{n,p} \\ 0_n & M & 0_{n,p} \\ 0_{p,n} & 0_{p,n} & 0_p \end{bmatrix} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} + \begin{bmatrix} 0_n & -I_n & 0_{n,p} \\ K & D & -C \\ C^T & 0_{p,n} & 0_p \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} 0_n \\ I_n \\ 0_{p,n} \end{bmatrix} f(t), \quad (6.2)$$

where a single subscript  $n$  or  $p$  indicate  $n \times n$  or  $p \times p$  square matrices respectively, and double subscripts  $n, p$  or  $p, n$  indicate  $n \times p$  or  $p \times n$  rectangular matrices respectively.

We let

$$\begin{aligned} E &= \begin{bmatrix} I_n & 0_n & 0_{n,p} \\ 0_n & M & 0_{n,p} \\ 0_{p,n} & 0_{p,n} & 0_p \end{bmatrix} \\ B &= \begin{bmatrix} 0_n & -I_n & 0_{n,p} \\ K & D & -C \\ C^T & 0_{p,n} & 0_p \end{bmatrix} \\ x(t) &= \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \\ q(t) &= \begin{bmatrix} 0_n \\ I_n \\ 0_{p,n} \end{bmatrix} f(t). \end{aligned}$$

This yields the following DAE

$$E\dot{x}(t) + Bx(t) = q(t). \quad (6.3)$$

We now present the first result.

**Proposition 6.1.** *Let the matrix  $C$  have full rank. Then the DAE (6.1) is solvable if and only if  $\det(\lambda M + D + \frac{1}{\lambda}K) \neq 0$  for some  $\lambda \neq 0$ .*

*Proof.* Theorem 2.3.1 from Brenan, Campbell and Petzold [3] states that the DAE (6.3) is solvable if and only if  $\det(\lambda E + B) \neq 0$ , i.e.,  $\det(\lambda E + B)$  is not identically zero as a function of  $\lambda$ . We now apply this solvability condition to our specific matrices involved in the DAE (6.3). We get the following block-partitioned matrix result for the matrix pencil  $(\lambda E + B)$ ,

$$\lambda E + B = \begin{bmatrix} \lambda I_n & -I_n & 0_{n,p} \\ K & \lambda M + D & -C \\ C^T & 0_{p,n} & 0_p \end{bmatrix}. \quad (6.4)$$

We now drop the  $p$  and  $n$  subscripts. The size of an individual block matrix can be determined by its position in the larger block-partitioned matrix. Specifically, the partitioned matrix has  $n$  rows for each block matrix in its first two partitioned rows and  $p$  rows for each block matrix in its third partitioned row. The same applies respectively for the columns of each block matrix based on the partitioned column in which it is located in the larger matrix.

We will use two properties of determinants: (1)  $\det F \det G = \det(FG)$ , and (2) the determinant of an upper or lower triangular block matrix is the product of the determinants of the matrices on the diagonal. Using block matrix Gaussian elimination to perform row reduction on (6.4) yields: (We assume  $\lambda \neq 0$ .)

$$\begin{aligned} \det \begin{bmatrix} \lambda I & -I & 0 \\ K & \lambda M + D & -C \\ C^T & 0 & 0 \end{bmatrix} &= \det \underbrace{\begin{bmatrix} I & 0 & 0 \\ -\frac{1}{\lambda}K & I & 0 \\ 0 & 0 & I \end{bmatrix}}_{=\det I \det I \det I=1} \det \begin{bmatrix} \lambda I & -I & 0 \\ K & \lambda M + D & -C \\ C^T & 0 & 0 \end{bmatrix} \\ &= \det \begin{bmatrix} \lambda I & -I & 0 \\ 0 & \lambda M + D + \frac{1}{\lambda}K & -C \\ C^T & 0 & 0 \end{bmatrix} \\ &= \det \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -\frac{1}{\lambda}C^T & 0 & I \end{bmatrix} \det \begin{bmatrix} \lambda I & -I & 0 \\ 0 & \lambda M + D + \frac{1}{\lambda}K & -C \\ C^T & 0 & 0 \end{bmatrix} \\ &= \det \begin{bmatrix} \lambda I & -I & 0 \\ 0 & \lambda M + D + \frac{1}{\lambda}K & -C \\ 0 & \frac{1}{\lambda}C^T & 0 \end{bmatrix} \\ &= \lambda^n \det \begin{bmatrix} \lambda M + D + \frac{1}{\lambda}K & -C \\ \frac{1}{\lambda}C^T & 0 \end{bmatrix}. \end{aligned}$$

At this point, note that the element  $\lambda M + D + \frac{1}{\lambda}K$  is a square matrix. The other two nonzero matrices contained in the last determinant matrix are rectangular. In order to proceed we now assume that the matrix  $\lambda M + D + \frac{1}{\lambda}K$  is invertible, i.e.,

$\det(\lambda M + D + \frac{1}{\lambda}K) \neq 0$ . Then,

$$\begin{aligned} \det \begin{bmatrix} \lambda M + D + \frac{1}{\lambda}K & -C \\ \frac{1}{\lambda}C^T & 0 \end{bmatrix} &= \det \begin{bmatrix} I & 0 \\ -\frac{1}{\lambda}C^T(\lambda M + D + \frac{1}{\lambda}K)^{-1} & I \end{bmatrix} \det \begin{bmatrix} \lambda M + D + \frac{1}{\lambda}K & -C \\ \frac{1}{\lambda}C^T & 0 \end{bmatrix} \\ &= \det \begin{bmatrix} \lambda M + D + \frac{1}{\lambda}K & -C \\ 0 & \frac{1}{\lambda}C^T(\lambda M + D + \frac{1}{\lambda}K)^{-1}C \end{bmatrix} \\ &= \det(\lambda M + D + \frac{1}{\lambda}K) \det[\frac{1}{\lambda}C^T(\lambda M + D + \frac{1}{\lambda}K)^{-1}C]. \end{aligned}$$

Therefore, since  $C$  (and  $C^T$ ) is full rank, the requirement that  $\det(\lambda E + B) \neq 0$  results in the condition that  $\det(\lambda M + D + \frac{1}{\lambda}K) \neq 0$ .

Thus, the DAE (6.1) is solvable if and only if  $\det(\lambda M + D + \frac{1}{\lambda}K) \neq 0$  for some  $\lambda \neq 0$ .  $\square$

So far we have only used the fact that  $C$  is full rank. If we add the additional condition that  $M$  is nonsingular, then we can obtain an additional proposition. However, before proceeding we will need the following lemma from functional analysis (see for example Taylor and Lay [38], Theorem IV.1.5.).

**Lemma 6.1.** *Let  $X$  be a Banach space and let  $L(X)$  be a Banach algebra of bounded linear operators which map  $X$  into itself. By definition, an operator  $M$  is invertible in  $L(X)$  if  $M$  is bijective and  $M^{-1} \in L(X)$ , i.e.,  $M^{-1}$  is bounded. If  $M$  is invertible in  $L(X)$ ,  $N \in L(X)$ , and  $\|M - N\| < \frac{1}{\|M^{-1}\|}$ , then  $N$  is invertible in  $L(X)$ .*

**Proposition 6.2.** *If  $C$  is full rank and  $M$  is nonsingular, then the DAE (6.1) is solvable.*

*Proof.* From Proposition 6.1 assuming  $C$  is full rank, we obtained that the DAE (6.1) is solvable if the matrix  $\lambda M + D + \frac{1}{\lambda}K$  is invertible for some  $\lambda \neq 0$ . We show that this matrix is invertible for  $\lambda$  satisfying  $|\lambda| > \|M^{-1}\|(\|D\| + \|K\|)$  or  $|\lambda| > 1$  whichever is larger. In the following we assume  $|\lambda| > 1$ ,

$$\begin{aligned} \|M^{-1}\|(\|D\| + \|K\|) &< |\lambda| \\ \|D\| + \|K\| &< \frac{1}{\|(\lambda M)^{-1}\|} \\ \|D + \frac{1}{\lambda}K\| &\leq \|D\| + \frac{1}{|\lambda|}\|K\| < \|D\| + \|K\| < \frac{1}{\|(\lambda M)^{-1}\|} \\ \|-D - \frac{1}{\lambda}K\| &< \frac{1}{\|(\lambda M)^{-1}\|} \\ \|\lambda M - \lambda M - D - \frac{1}{\lambda}K\| &< \frac{1}{\|(\lambda M)^{-1}\|} \\ \|\lambda M - (\lambda M + D + \frac{1}{\lambda}K)\| &< \frac{1}{\|(\lambda M)^{-1}\|}. \end{aligned}$$



Since  $M$  is invertible,  $\lambda M$  is also invertible for  $\lambda \neq 0$ . Now by Lemma 6.1,  $\lambda M + D + \frac{1}{\lambda}K$  is invertible. Thus for  $M$  invertible, we have  $\det(\lambda M + D + \frac{1}{\lambda}K) \neq 0$  for some  $\lambda$ , specifically for any  $|\lambda| > \max \{1, \|M^{-1}\|(\|D\| + \|K\|)\}$ , or equivalently,  $|\lambda| > \max \{1, \frac{(\|D\| + \|K\|)}{\|M\|}\}$ .  $\square$

## 6.2 Incorporating the Algebraic Constraints into an Explicit ODE

We now look at a different formulation of the DAE (6.1) by using the constraint equation  $C^T z(t) = 0$  to solve for  $F(t)$  and form an explicit ordinary differential equation (ODE). Based on our results from Section 6.1 and given that  $C$  is full rank and  $M > 0$  is symmetric, we expect this method to be feasible since we know this particular DAE is solvable.

We start with the second-order ODE equation. Since  $M > 0$  is invertible, we multiply by  $M^{-1}$  and then by the transpose  $C^T$ . We have

$$\begin{aligned} M\ddot{z}(t) + D\dot{z}(t) + Kz(t) &= CF(t) + f(t) \\ \ddot{z}(t) + M^{-1}D\dot{z}(t) + M^{-1}Kz(t) &= M^{-1}CF(t) + M^{-1}f(t) \\ C^T\ddot{z}(t) + C^TM^{-1}D\dot{z}(t) + C^TM^{-1}Kz(t) &= C^TM^{-1}CF(t) + C^TM^{-1}f(t). \end{aligned}$$

We differentiate the constraint equation twice to obtain  $C^T\ddot{z}(t) = 0$  and apply to the last equation above which yields

$$\begin{aligned} C^TM^{-1}D\dot{z}(t) + C^TM^{-1}Kz(t) &= C^TM^{-1}CF(t) + C^TM^{-1}f(t) \\ C^TM^{-1}CF(t) &= C^TM^{-1}D\dot{z}(t) + C^TM^{-1}Kz(t) - C^TM^{-1}f(t) \\ F(t) &= (C^TM^{-1}C)^{-1}C^TM^{-1}D\dot{z}(t) + (C^TM^{-1}C)^{-1}C^TM^{-1}Kz(t) \\ &\quad - (C^TM^{-1}C)^{-1}C^TM^{-1}f(t), \end{aligned}$$

where we have used the fact  $C$  is full rank in order to justify the existence of  $(C^TM^{-1}C)^{-1}$ . We then substitute this result for  $F(t)$  back into the original second-order ODE:

$$\begin{aligned} M\ddot{z}(t) + D\dot{z}(t) + Kz(t) - C(C^TM^{-1}C)^{-1}C^TM^{-1}D\dot{z}(t) - C(C^TM^{-1}C)^{-1}C^TM^{-1}Kz(t) \\ = f(t) - C(C^TM^{-1}C)^{-1}C^TM^{-1}f(t) \\ M\ddot{z}(t) + [I - C(C^TM^{-1}C)^{-1}C^TM^{-1}]D\dot{z}(t) + [I - C(C^TM^{-1}C)^{-1}C^TM^{-1}]Kz(t) \\ = [I - C(C^TM^{-1}C)^{-1}C^TM^{-1}]f(t). \end{aligned}$$

We define the matrix  $Q \in \mathbb{R}^{n \times n}$  as follows

$$Q = C(C^TM^{-1}C)^{-1}C^TM^{-1}. \quad (6.5)$$

We then let  $P := I - Q \in \mathbb{R}^{n \times n}$  which then leads to the second-order ODE

$$M\ddot{z}(t) + PD\dot{z}(t) + PKz(t) = Pf(t). \quad (6.6)$$

We now convert this to a first-order ODE where  $x_1(t) = z(t)$  and  $x_2(t) = \dot{z}(t) = \dot{x}_1(t)$ ,

$$\begin{bmatrix} I_n & 0_n \\ 0_n & M \end{bmatrix} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} + \begin{bmatrix} 0_n & -I_n \\ PK & PD \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 0_{n,1} \\ Pf(t) \end{bmatrix},$$

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & M^{-1} \end{bmatrix} \begin{bmatrix} 0 & -I \\ PK & PD \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & M^{-1} \end{bmatrix} \begin{bmatrix} 0 \\ Pf(t) \end{bmatrix},$$

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} + \begin{bmatrix} 0 & -I \\ M^{-1}PK & M^{-1}PD \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 0 \\ M^{-1}Pf(t) \end{bmatrix}.$$

We note that the initial conditions must satisfy  $C^T x_1(0) = 0$  and  $C^T x_2(0) = 0$  to ensure we start out on the constraint manifold. Thus, our final system becomes a regular initial value problem ODE:

$$\left\{ \begin{array}{l} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} + \begin{bmatrix} 0 & -I \\ M^{-1}PK & M^{-1}PD \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 0 \\ M^{-1}Pf(t) \end{bmatrix} \\ \begin{bmatrix} C^T & 0 \\ 0 & C^T \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{array} \right. \quad (6.7)$$

with a constraint imposed on the initial data, i.e., we require consistent initial data.

*Remark.* Not all solutions of the explicit ODE are solutions of the original DAE. However, by imposing *consistent initial conditions* to ensure we start out in the constraint manifold,  $\{z \in \mathbb{R}^n : C^T z(t) = 0\}$ , will ensure that the solution of the ODE and the solution of the DAE are equivalent. However, when solving the explicit ODE numerically, one must be aware of the issue of numerical drift when  $z(t)$  may drift off of the constraint manifold.

Letting

$$\begin{aligned} x(t) &= \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \\ A &= \begin{bmatrix} 0 & I \\ -M^{-1}PK & -M^{-1}PD \end{bmatrix}, \\ q(t) &= \begin{bmatrix} 0 \\ M^{-1}Pf(t) \end{bmatrix}, \\ G &= \begin{bmatrix} C^T & 0 \\ 0 & C^T \end{bmatrix}, \end{aligned}$$

we have

$$\begin{cases} \dot{x}(t) = Ax(t) + q(t) \\ Gx(0) = 0. \end{cases} \quad (6.8)$$

*Remark.* Notice that regardless of whether we started with  $C^T \dot{z}(t) = 0$  or  $C^T z(t) = 0$  we would still get the same explicit ODE (6.7); the only difference is whether we had to differentiate the constraint twice or just once before solving for the algebraic variable  $F(t)$ . However, if we started with  $C^T \dot{z}(t) = 0$ , we would need to be careful that the initial condition  $C^T z(0) = 0$  makes sense and is consistent. Additionally, even if the DAE (6.1) were solvable it might not always be possible to form an explicit ODE. This would be true in particular for implicit or nonlinear DAEs.

We now investigate the matrices  $Q$  and  $P$ . As noted above,  $Q = C(C^T M^{-1} C)^{-1} C^T M^{-1}$  and  $P = I - Q$ .  $Q$  is idempotent since

$$\begin{aligned} Q^2 &= C(C^T M^{-1} C)^{-1} (C^T M^{-1} C) (C^T M^{-1} C)^{-1} C^T M^{-1} \\ &= C(C^T M^{-1} C)^{-1} C^T M^{-1} \\ &= Q. \end{aligned}$$

Thus, we also have  $P^2 = P$ .

Before we give details on what  $Q$  does, we first review information on projections from Householder, Section 1.3, [22]. Two vectors in  $\mathbb{R}^n$  are orthogonal if  $y^T x = 0$ . We can always decompose any vector  $z \in \mathbb{R}^n$  into two orthogonal components  $z = x + y$  such that  $y^T x = 0$ . If  $Qz = y$  and  $(I - Q)z = Pz = x$ , then  $Q$  is the orthogonal projection onto  $\text{im } Q$  and  $P$  is the orthogonal projection onto  $(\text{im } Q)^\perp$  with  $Q^2 = Q$ ,  $Q^T = Q$ , and  $PQ = QP = 0$ . Thus, we have  $(Qz)^T (Pz) = z^T Q^T Pz = z^T QPz = 0$  for all  $z \in \mathbb{R}^n$ , or  $x$  and  $y$  are orthogonal for  $z = x + y$  where  $x = Pz$  and  $y = Qz$ .

Likewise if a matrix  $C \in \mathbb{R}^{n \times p}$  has full rank  $p$ ,  $p \leq n$ , and if the vector  $Cx$  is the orthogonal projection of a vector  $z \in \mathbb{R}^n$  onto the space spanned by the columns of  $C$ , then  $C^T(z - Cx) = 0$  since  $\ker C^T = (\text{im } C)^\perp$ . Solving for  $x$  yields

$$\begin{aligned} C^T(z - Cx) &= C^T z - C^T Cx = 0 \\ C^T Cx &= C^T z \\ x &= (C^T C)^{-1} C^T z, \end{aligned}$$

where  $(C^T C)^{-1}$  exists because  $C$  has full column rank. Note that  $x \in \mathbb{R}^p$  with  $\dim(\text{im } C) = p$ . Hence  $\tilde{Q}z = Cx = C(C^T C)^{-1} C^T z$  or  $\tilde{Q} = C(C^T C)^{-1} C^T$  is the orthogonal projection onto  $\text{im } C$ . It is easy to check  $\tilde{Q}^2 = \tilde{Q}$ ,  $\tilde{Q}^T = \tilde{Q}$ . Furthermore,  $\tilde{P} = I - \tilde{Q}$  is then the projection onto  $(\text{im } C)^\perp = \ker C^T$ .

We next generalize the above result. We say that, given a symmetric positive definite matrix  $G$ , two vectors  $x$  and  $y$  are *orthogonal with respect to  $G$*  if  $y^T Gx = 0$ . To

give some understanding to this concept, a symmetric positive definite matrix  $G$  can always be factored by a Cholesky factorization such that  $G = B^T B$ . Then, we can interpret  $y^T G x = y^T B^T B x = (B y)^T (B x) = 0$  as the ordinary scalar product between two orthogonal vectors  $B y$  and  $B x$  where  $x$  and  $y$  can be considered basis vectors in an oblique coordinate system and then  $B x$  and  $B y$  are the respective basis vectors after transformation into an orthogonal coordinate system.

We now repeat the derivation of  $Q$  as above except we generalize so that the projection is orthogonal with respect to  $G$ . In our case, we let  $G = M^{-1}$ . Thus, if  $M^{-1}$  is symmetric positive definite, if  $C$  has full column rank  $p$ , and if we assume  $C x$  is the orthogonal projection with respect to  $M^{-1}$  of a vector  $z \in \mathbb{R}^n$  onto the space spanned by the columns of  $C$ , then  $C^T M^{-1}(z - C x) = 0$  since now  $\ker C^T M^{-1} = (\text{im } C)^\perp$ . Solving again for  $x$  yields

$$\begin{aligned} C^T M^{-1}(z - C x) &= C^T M^{-1} z - C^T M^{-1} C x = 0 \\ C^T M^{-1} C x &= C^T M^{-1} z \\ x &= (C^T M^{-1} C)^{-1} C^T M^{-1} z. \end{aligned}$$

Hence,  $Q z = C x = C(C^T M^{-1} C)^{-1} C^T M^{-1} z$  or  $Q = C(C^T M^{-1} C)^{-1} C^T M^{-1}$ . We then define  $P = I - Q$ . Thus,  $Q$  is the orthogonal projection with respect to  $M^{-1}$  onto  $\text{im } C$ . We note for information only that in this generalized projection case we do not have  $Q^T = Q$ , i.e.,  $Q$  is not symmetric. Since  $(\text{im } C)^\perp = \ker C^T$ ,  $P$  is the orthogonal projection with respect to  $M^{-1}$  onto  $\ker C^T$ .

Another way to express this result is that  $C^T M^{-1} P z = 0$  for any  $z \in \mathbb{R}^n$ . Hence,  $P$  can also be viewed as a projection onto  $\ker(C^T M^{-1})$ . Therefore,  $C^T M^{-1} P K x(t) = 0$ ,  $C^T M^{-1} P D \dot{x}(t) = 0$ , and  $C^T M^{-1} P f(t) = 0$ . Thus, multiplying the second-order ODE (6.6) above by  $C^T M^{-1}$ , gives  $C^T \ddot{z}(t) = 0$  and shows that we satisfy our constraint equation. Finally, in equation (6.7) above,  $M^{-1} P$  can be interpreted as a mapping onto  $\ker C^T$ , i.e., onto the constraint manifold  $\{z(t) \in \mathbb{R}^n : C^T z(t) = 0\}$ .

*Remark.* In our original two beam and joint problem,  $M$  was defined as a positive definite symmetric mass matrix. Hence, we can assume that  $M^{-1}$  is also positive definite symmetric and our interpretation of the projection  $P$  is valid.

### 6.3 Hessenberg Form of DAEs

Before we look at the Hessenberg form of DAEs, we first recall the definition of the index of a DAE.

**Definition 6.1.** For the DAE (6.3), the *index* can be defined as the number of iterations of the following process: (i) Perform coordinate changes to rewrite the DAE with explicit

algebraic constraints, and then (ii) Differentiate the algebraic constraints. Repeat until the system is reduced to an explicit ODE.

*Remark.* In the last section we formed an explicit ODE by differentiating the DAE (6.1) constraints twice and substituting back in to solve for the unknown algebraic variables  $F(T)$ . Based on this derivation of the explicit ODE and the definition of index, one may think that the index for DAE (6.1) is two.

Since the DAE (6.1) has a  $3 \times 3$  block matrix when put into a first order system (6.2), let us look at the standard Hessenberg size-3 form of a DAE.

$$\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & 0 \\ 0 & B_{32} & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = q(t). \quad (6.9)$$

**Definition 6.2.** A DAE is said to be in *Hessenberg size-3 form* if it has the standard form above (6.9) and meets the condition that the product  $B_{32}B_{21}B_{13}$  is square and nonsingular.

We compare the DAE (6.2) with the Hessenberg size-3 form by multiplying the second row of (6.2) by  $M^{-1}$ ,

$$\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} + \begin{bmatrix} 0 & -I & 0 \\ M^{-1}K & M^{-1}D & -M^{-1}C \\ C^T & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} 0 \\ M^{-1}f(t) \\ 0 \end{bmatrix}.$$

In this case, the product  $B_{32}B_{21}B_{13} = 0$ . Before we say this is not a Hessenberg size-3 DAE, we notice that by reordering the  $x_1$  and  $x_2$  variables we can maintain our structure in the left matrix in front of the differentiated variables. Then, the effect of reordering the variables on the  $B$  matrix is to swap the first two rows and then switch the first two columns. We then get the following reordered system:

$$\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_2(t) \\ \dot{x}_1(t) \\ \dot{x}_3(t) \end{bmatrix} + \begin{bmatrix} M^{-1}D & M^{-1}K & -M^{-1}C \\ -I & 0 & 0 \\ 0 & C^T & 0 \end{bmatrix} \begin{bmatrix} x_2(t) \\ x_1(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} M^{-1}f(t) \\ 0 \\ 0 \end{bmatrix}. \quad (6.10)$$

Now the product  $B_{32}B_{21}B_{13} = C^T M^{-1}C$  is square and invertible since  $C$  is full rank. Hence, our original DAE (6.1) is a Hessenberg size-3 DAE. Why is this significant?

From Brenan, Campbell, and Petzold [3], a Hessenberg size- $i$  DAE is solvable and has index  $i$ . Hence, our original DAE is solvable and has index three. The fact that it is solvable agrees with our results from section 6.1.

*Remark.* In section 6.2 we assumed that the index of this DAE was two, based on differentiating the constraint equation twice. Hence, differentiating may only give a minimum index number and may not be an accurate indicator of the index.

We now look at a modified version of (6.1). Specifically, we will look at the differentiated constraint  $C^T \dot{z}(t) = 0$  vice  $C^T z(t) = 0$ . Our modified system becomes

$$\begin{cases} M\ddot{z}(t) + D\dot{z}(t) + Kz(t) - CF(t) = f(t), \\ C^T \dot{z}(t) = 0. \end{cases} \quad (6.11)$$

When put into first order form we get:

$$\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} + \begin{bmatrix} 0 & -I & 0 \\ M^{-1}K & M^{-1}D & -M^{-1}C \\ 0 & C^T & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} 0 \\ M^{-1}f(t) \\ 0 \end{bmatrix}.$$

Here we have multiplied the second row by  $M^{-1}$  to try to put the system into Hessenberg form. The only difference between the original DAE and the modified DAE is that the  $C^T$  term is in the bottom center element of matrix  $B$  vice in the bottom left element.

We note immediately that the product  $B_{32}B_{21}B_{13} = 0$ . If we reorder the  $x_1$  and  $x_2$  variables as before the effect on the  $B$  matrix is to swap the first two rows and then switch the first two columns. This will move the  $C^T$  term from the bottom center element back to the bottom left element. Then,  $B_{32} = 0$  and hence the product  $B_{32}B_{21}B_{13} = 0$ . Thus, this modified DAE is not a Hessenberg size-3 DAE.

However, if we partition the  $B$  matrix for this modified DAE, we have

$$B = \begin{bmatrix} 0 & -I & \vdots & 0 \\ M^{-1}K & M^{-1}D & \vdots & -M^{-1}C \\ \cdots & \cdots & \cdots & \cdots \\ 0 & C^T & \vdots & 0 \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & 0 \end{bmatrix},$$

where the product  $B_{21}B_{12} = -C^T M^{-1}C$  is square and nonsingular.

Thus, we get a Hessenberg size-2 DAE:

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & 0 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} q_1(t) \\ q_2(t) \end{bmatrix}, \quad (6.12)$$

where  $y_1(t) = [x_1(t), x_2(t)]^T$ ,  $y_2(t) = x_3(t)$ ,  $q_1(t) = [0 \ M^{-1} f(t)]^T$ ,  $q_2(t) = 0$ . Thus, this modified DAE with the constraint differentiated once results in a solvable DAE and has index two.

*Remark.* It may not always be obvious that a DAE can be put into Hessenberg form. It may take some work. But a word of warning, not all DAEs can be put into Hessenberg form. Hence, if a DAE cannot be put into Hessenberg size-3 or size-2 does not mean it is not an index-3 or index-2 DAE respectively. However, if the DAE can be put into a Hessenberg form, it immediately tells us that the DAE is solvable and has index of " $i$ " associated with size- $i$  Hessenberg form.

*Remark.* We note that differentiating the constraint equation once reduced the index of the DAE by one as expected.

# Chapter 7

## Conclusions

### 7.1 Summary and Contributions

Using the detailed framework and background developed by März [33] we have extended the theory from ordinary DAEs to abstract DAEs involving systems of partial differential equations and hybrid systems. We focused our initial efforts on index-1 DAEs and those of semi-explicit form. While one could easily use a more direct approach for the index-1 semi-explicit DAE, this projection and subspace approach has much more potential for those problems that do not exactly fit the semi-explicit index-1 form. Hence, our thorough analysis of this method and the thorough example application of a structural dynamics problem will be helpful in continuing the theory beyond its current bounds as well as being useful for applications that may be index-1 abstract DAEs but do not allow use of a direct method.

We have also started a theory for abstract Hessenberg DAEs in the index-2 abstract DAE case. The theory for index-2 is much more difficult as seen. In the index-1 case, the  $P_0$  and  $Q_0$  projectors only depend on  $\mathcal{E}$  which in this case only included the identity operator and zero operators. When moving into the index-2 case, even for the simplified semi-explicit DAE, the projectors  $P_1$  and  $Q_1$  become much more complex. They now depend on the operators making up the matrix operator  $\mathcal{B}$ . We have to worry about whether these index-2 projectors now “commute” with the various  $\mathcal{B}_{ij}$  operators, i.e., whether the images of the projectors are in the domains of the  $\mathcal{B}_{ij}$  operators, and whether we lose infinitesimal generator and dissipative operator properties of the  $\mathcal{B}_{ij}$  operators when composed with these index-2 projectors. This was one of the issues we attempted to address for the abstract Hessenberg DAE and the associated projector  $H$ . In an attempt to generalize the problem we also defined a concept of an operator domain that is *uniformly dense* with respect to an orthogonal projector  $P$  on a Hilbert space.



In the final chapter, we addressed some of the issues that arise when working with the discretized DAE. Specifically, when an abstract DAE is discretized using a finite element method for the spatial variables, we may end up with an ordinary DAE. What form of the now discretized DAE do we want to work with? Can we work with the reduced index form of the DAE by differentiating the constraints? What is the difference between converting the discretized DAE to its underlying or explicit ODE? What are the drawbacks of using the explicit ODE form? We explored some of these issues as well as converting a DAE to Hessenberg form. However, we do not provide any answers to these questions at this point. This exploration primarily whets the appetite for more work yet to be done in this area. Moreover, a comparison between the form of the underlying explicit ODE and that of the inherent regular ODE defined by März is interesting. There may be a way to efficiently perform a projection at each step, or even more likely after a series of steps, of the numerical approximation to help prevent drift off of the constraint manifold. Determining how long one can proceed between correcting projection steps would be critical. This idea could lead to a kind of “corrector” step which could be added to existing algorithms.

## 7.2 Future Work

There is much work to be done in this area. Some related research is ongoing in the area of PDAEs. However, we have only begun to scratch the surface of what needs to be done for the broader scope of abstract DAEs. For example, in the area of well-posedness more work still needs to be done with respect to continuous dependence on initial data. Does this concept even apply? We only looked at existence and uniqueness of solutions. Additionally, we primarily focused on semigroup methods for determining existence and uniqueness of solutions to the AODE. There is much more literature on standard elliptic and parabolic PDE theory. Thus, determining what it means to have a weak or mild solution of a DAE would be a useful step forward.

Much more effort is needed for the abstract index-2 case. We only addressed the abstract Hessenberg size-2 problem. It would be useful to attack a more general index-2 version. The other thing to factor in are the applications. It would be helpful to first have some good abstract index-2 problems to work with to help guide possible solutions. We also only addressed the linear time-invariant case where our  $\mathcal{E}$  and  $\mathcal{B}$  operators were constant and not a function of time. It would be useful to extend the abstract DAE theory to the linear time-varying problem as well. Here the methodology developed by März should also be helpful. Ordinary DAEs are already being used in the control literature. An extension to abstract DAEs could be made for the infinite-dimensional control problem or distributed parameter systems.

There has already been some work in the area of numerical approximation to PDAEs. As discussed above in Section 7.1 there is some potential for modifying algorithms to include correction steps depending on the type of algorithm. Iterative methods in particular are desired for the large matrices that result from discretization of PDEs. Stability, accuracy and efficiency need to be addressed and resolved for the different algorithms and methods used for numerically solving abstract DAEs. A general theory of index reduction for discretized abstract DAEs would prove to be very useful and allow more problems to be solved using existing algorithms.

As one can see there are many openings remaining for productive research in the area of abstract DAEs. What are you waiting for? I for one will continue from where this dissertation ends.

# Bibliography

- [1] U. M. Ascher and L. R. Petzold. *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. SIAM, Philadelphia, 1998.
- [2] K. Balla and R. März. A unified approach to linear differential algebraic equations and their adjoint equations. *Z. Anal. Anwendungen*, 21(3):783-802, 2002.
- [3] K. E. Brenan, S. L. Campbell and L. R. Petzold. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. SIAM, Philadelphia, 1996.
- [4] Y. Boyarintsev. *Methods of Solving Singular Systems of Ordinary Differential Equations*. John Wiley & Sons, New York, 1992.
- [5] S. L. Campbell. *Singular Systems of Differential Equations*. Pitman, Boston, 1980.
- [6] S. L. Campbell. *Singular Systems of Differential Equations II*. Pitman, Boston, 1982.
- [7] S. L. Campbell and W. Marszalek. Index of infinite dimensional differential algebraic equations, *Math. Comp. Model. Dynam. Syst.* 5(1):18–42, 1999.
- [8] R. W. Carroll and R. E. Showalter. *Singular and Degenerate Cauchy Problems*. Academic Press, New York, 1976.
- [9] E. M. Cliff and T. L. Herdman. Well-Posedness for a Model of Radiant Energy Exchange, Interdisciplinary Center for Mathematics Report 99-05-01, Virginia Tech, Blacksburg, VA, 1999.
- [10] R. F. Curtain and A. J. Pritchard. *Infinite Dimensional Linear Systems Theory*. Springer-Verlag, New York, 1978.
- [11] N. Dunford and J. T. Schwartz. *Linear Operators Part I: General Theory*. John Wiley & Sons, New York, 1958.
- [12] K.-J. Engel and R. Nagel. *One-Parameter Semigroups for Linear Evolution Equations*. Springer-Verlag, New York, 2000.

- [13] H. O. Fattorini. *The Cauchy Problem*. Encyclopedia of Mathematics and Its Applications, Vol. 18, Addison-Wesley, Reading, 1983.
- [14] H. O. Fattorini. *Second Order Linear Differential Equations in Banach Spaces*. Elsevier, Amsterdam, 1985.
- [15] A. Favini and A. Yagi. *Degenerate Differential Equations in Banach Spaces*. Marcel Dekker, New York, 1999.
- [16] F. R. Gantmacher. *The Theory of Matrices*. AMS, Providence, 1998.
- [17] J. A. Goldstein. *Semigroups of Linear Operators and Applications*. Oxford University Press, New York, 1985.
- [18] G. H. Golub and C. F. Van Loan. *Matrix Computations*. John Hopkins University Press, Baltimore, 1996.
- [19] E. Griepentrog and R. März. *Differential-Algebraic Equations and Their Numerical Treatment*. Teubner, Leipzig, 1986.
- [20] E. Hairer, S. P. Norsett and G. Wanner. *Solving Ordinary Differential Equations II*. Springer-Verlag, New York, 1993.
- [21] E. Hille and R. S. Phillips. *Functional Analysis and Semi-groups*. AMS, Providence, 1957.
- [22] A. S. Householder. *The Theory of Matrices in Numerical Analysis*. Dover, New York, 1975.
- [23] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, New York, 1980.
- [24] M. V. Keldysh. On eigenvalues and eigenfunctions of some classes of nonselfadjoint equations, Dokl. Akad. Nauk. USSR, 77(1):11–14, 1951.
- [25] S. G. Krein. *Linear Differential Equations in Banach Space*. AMS, Providence, 1971.
- [26] E. Kreyzsig. *Introductory Functional Analysis with Applications*. John Wiley & Sons, New York, 1978.
- [27] G. E. Ladas and V. Lakshmikantham. *Differential Equations in Abstract Spaces*. Academic Press, New York, 1972.
- [28] P. D. Lax. *Functional Analysis*. John Wiley & Sons, New York, 2002.

- [29] R. Lamour, R. März, and C. Tischendorf. PDAEs and Further Mixed Systems as Abstract Differential Algebraic Systems, Preprint 2001-11, Institut für Mathematik, Humboldt-Universität Berlin, 2001.  
<http://www.mathematik.hu-berlin.de/publ/pre/2001/p-list-01.html>
- [30] Z. Liu and S. Zheng. *Semigroups associated with dissipative systems*, Chapman and Hall/CRC, Boca Raton, 1999.
- [31] W. Lucht, K. Strehmel and C. Eichler-Liebenow. Indexes and Special Discretization Methods for Linear Partial Differential Algebraic Equations, BIT, 39(3):484–512, 1999.
- [32] A. S. Markus. *Introduction to the Spectral Theory of Polynomial Operator Pencils*. AMS, Providence, 1988.
- [33] R. März. Differential algebraic systems anew. Appl. Numer. Math., 42(1-3):315–335, 2002.
- [34] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer-Verlag, New York, 1983.
- [35] L. R. Petzold. Differential/Algebraic Equations are not ODEs, SIAM J. Sci. Statist. Comput., 3(3):367–384, 1982.
- [36] P. J. Rabier and W. C. Rheinboldt. *Nonholonomic Motion of Rigid Mechanical Systems From a DAE Viewpoint* SIAM, Philadelphia, 2000.
- [37] G. J. Reid, P. Lin and A. D. Wittkopf. Differential Elimination - Completion Algorithms for DAE and PDAE, Stud. Appl. Math., 106(1):1–45, 2001.
- [38] A. E. Taylor and D. C. Lay. *Introduction to Functional Analysis*. John Wiley & Sons, New York, second edition, 1980.
- [39] S. Yakubov and Y. Yakubov. *Differential-Operator Equations: Ordinary and Partial Differential Equations*. Chapman and Hall/CRC, Boca Raton, 2000.
- [40] W. Walter. *Ordinary Differential Equations*. Springer-Verlag, New York, 1998.
- [41] S. Zaidman. *Abstract differential equations*. Pitman, London, 1979.
- [42] S. Zaidman. *Topics in abstract differential equations*. Longman Scientific & Technical, Harlow, 1994.

# Vita

## Mark A. Pierson

I graduated from the University of California, Davis with a Bachelor of Arts degree in German and was then commissioned in the U.S. Navy through Officer Candidate School. I had also taken three years of undergraduate physics and mathematics which is why I was accepted.

Following nuclear power training and submarine school, I reported to USS Simon Bolivar (SSBN 641)(Gold) at Portsmouth Naval Shipyard in Kittery, Maine. The ship completed demonstration and shakedown operations and then conducted four strategic deterrent patrols from Kings Bay, Georgia. During this tour I served as Interior Communications Officer, Electrical Officer, Reactor Controls Officer and Assistant Engineer. I also qualified in submarines entitling me to wear the submarine dolphin emblem. After a tour of almost three and a half years, I was transferred to the submarine repair ship USS Canopus (AS 34) in Charleston, South Carolina where I served as Radiological Controls Officer for two and a half years and completed qualification as Surface Warfare Officer. Upon completion of the Submarine Officer's Advanced Course for Department Heads, I was assigned as the Chief Engineer in USS Henry L. Stimson (SSBN 655)(Gold) at Charleston, South Carolina. While aboard Henry L. Stimson for two years, I completed four strategic deterrent patrols.

I then served for two years on the staff of the Director of Naval Reactors in Arlington, Virginia. This was the organization ADM Rickover founded to design and oversee the Naval nuclear propulsion program. While in Arlington, I pursued evening classes toward an MBA degree at George Washington University. Following my tour at Naval Reactors, I transferred to the California State University at Sacramento to continue my studies toward an MBA degree. While there I specialized in finance and artificial neural networks. I was also inducted into the Beta Gamma Sigma national honor society for business and management. Before I completed my thesis, I was transferred again to Honolulu, Hawaii as Executive Officer of USS Indianapolis (SSN 697), a Los Angeles class fast attack submarine. I never finished my thesis or MBA degree as a result. While on the USS Indianapolis, we conducted one Western Pacific deployment. After two years, I reported

to the Chief of Naval Operations staff for Submarine Warfare in Arlington, Virginia at the Pentagon. I served as the Requirements Officer for submarine maintenance.

After three and a half years in the Pentagon, I then had the pleasure of being assigned to the Office of Naval Research (ONR) as the Deputy Department Head for Engineering, Materials and Physical Sciences S & T. I thoroughly enjoyed this position for another three and a half years. While at ONR, I received my Master's degree in Mathematics from Virginia Tech's Falls Church campus through evening courses.

I then retired from the Navy at the rank of Commander following twenty-three years of service in June 2001. My military awards included the Meritorious Service Medal (two awards), the Navy Commendation Medal (three awards), the Navy Achievement Medal (two awards), the Navy Expeditionary Medal and the National Defense Medal.

While stationed at the Pentagon in Arlington, Virginia (now that I was no longer under water), I met my wife Rebekah Paulson at a Lutheran church. We were married in 1997. I also served as Scoutmaster and Assistant Scoutmaster of Arlington Boy Scout Troop 143 for four years.

Due to my new found love of mathematics while studying for a Master's degree, when I retired from the Navy I decided to pursue a doctoral degree full time in Blacksburg, Virginia. While at Virginia Tech I was inducted into the Pi Mu Epsilon mathematics honor society and the Phi Kappa Phi interdisciplinary honor society. I also served in the SIAM Student Chapter at Virginia Tech as its Vice-President. As Vice-President I coordinated a biweekly student research seminar given by graduate students including one outside guest speaker. While a graduate student in the Mathematics Department I received both Teaching Assistantships and a Research Assistantship. I taught various sections of calculus and differential equations. My goal is to some day be a professor of mathematics at a great university such as Virginia Tech. Now that I have finally finished my Ph.D. I recognize that my learning is just starting. There is so much more to learn in mathematics yet! I cannot wait!