Differential-Algebraic Equations Forum

DAE-F

René Lamour
Roswitha März
Caren Tischendorf

Differential-Algebraic Equations: A Projector Based Analysis

Springer

Differential-Algebraic Equations Forum

Differential-Algebraic Equations Forum

The series “Differential-Algebraic Equations Forum” is concerned with analytical, algebraic, control theoretic and numerical aspects of differential algebraic equations (DAEs) as well as their applications in science and engineering. It is aimed to contain survey and mathematically rigorous articles, research monographs and textbooks. Proposals are assigned to an Associate Editor, who recommends publication on the basis of a detailed and careful evaluation by at least two referees. The appraisals will be based on the substance and quality of the exposition.

For further volumes:
www.springer.com/series/11221

René Lamour • Roswitha März • Caren Tischendorf

Differential-Algebraic Equations: A Projector Based Analysis

Springer

René Lamour
Department of Mathematics
Humboldt-University of Berlin
Berlin, Germany

Roswitha März
Department of Mathematics
Humboldt-University of Berlin
Berlin, Germany

Caren Tischendorf
Mathematical Institute
University of Cologne
Cologne, Germany

ISBN 978-3-642-27554-8 ISBN 978-3-642-27555-5 (eBook)
DOI 10.1007/978-3-642-27555-5
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013930590

Mathematics Subject Classification (2010): 34A09, 34-02, 65L80, 65-02, 15A22, 65L05, 49K15, 34D05, 34G99

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword by the Editors

We are very pleased to write the Foreword of this book by René Lamour, Roswitha März, and Caren Tischendorf. This book appears as the first volume in the recently established series "FORUM DAEs"—a forum which aims to present different directions in the widely expanding field of differential-algebraic equations (DAEs).

Although the theory of DAEs can be traced back earlier, it was not until the 1960s that mathematicians and engineers started to study seriously various aspects of DAEs, such as computational issues, mathematical theory, and applications. DAEs have developed today, half a century later, into a discipline of their own within applied mathematics, with many relationships to mathematical disciplines such as algebra, functional analysis, numerical analysis, stochastics, and control theory, to mention but a few. There is an intrinsic mathematical interest in this field, but this development is also supported by extensive applications of DAEs in chemical, electrical and mechanical engineering, as well as in economics.

Roswitha März' group has been at the forefront of the development of the mathematical theory of DAEs since the early 1980s; her valuable contribution was to introduce—with a Russian functional analytic background—the method now known as the "projector approach" in DAEs. Over more than 30 years, Roswitha März established a well-known group within the DAE community, making many fundamental contributions. The projector approach has proven to be valuable for a huge class of problems related to DAEs, including the (numerical) analysis of models for dynamics of electrical circuits, mechanical multibody systems, optimal control problems, and infinite-dimensional differential-algebraic systems.

Broadly speaking, the results of the group have been collected in the present textbook, which comprises 30 years of development in DAEs from the viewpoint of projectors. It contains a rigorous and stand-alone introduction to the projector approach to DAEs. Beginning with the case of linear constant coefficient DAEs, this approach is then developed stepwise for more general types, such as linear DAEs with variable coefficients and nonlinear problems. A central concept in the theory of DAEs is the "index", which is, roughly speaking, a measure of the difficulty of (numerical) solution of a given DAE. Various index concepts exist in the theory of DAEs; and the one related to the projector approach is the “tractability index”. Analytical and numerical consequences of the tractability index are presented. In addition to the discussion of the analytical and numerical aspects of different classes of DAEs, this book places special emphasis on DAEs which are explicitly motivated by practice: The “functionality” of the tractability index is demonstrated by means of DAEs arising in models for the dynamics of electrical circuits, where the index has an explicit interpretation in terms of the topological structure of the interconnections of the circuit elements. Further applications and extensions of the projector approach to optimization problems with DAE constraints and even coupled systems of DAEs and partial differential equations (the so-called “PDAEs”) are presented.

If one distinguishes strictly between a textbook and a monograph, then we consider the present book to be the second available textbook on DAEs. Not only is it complementary to the other textbook in the mathematical treatment of DAEs, this book is more research-oriented than a tutorial introduction; novel and unpublished research results are presented. Nonetheless it contains a self-contained introduction to the projector approach. Also various relations and substantial cross-references to other approaches to DAEs are highlighted.

This book is a textbook on DAEs which gives a rigorous and detailed mathematical treatment of the subject; it also contains aspects of computations and applications. It is addressed to mathematicians and engineers working in this field, and it is accessible to students of mathematics after two years of study, and also certainly to lecturers and researchers. The mathematical treatment is complemented by many examples, illustrations and explanatory comments.

Ilmenau, Germany *Achim Ilchmann*
Hamburg, Germany *Timo Reis*
June 2012

Preface

We assume that differential-algebraic equations (DAEs) and their more abstract versions in infinite-dimensional spaces comprise *great potential for future mathematical modeling*. To an increasingly large extent, in applications, DAEs are automatically generated, often by coupling various subsystems with large dimensions, but *without manifested mathematically useful structures*. Providing tools to uncover and to monitor mathematical DAE structures is one of the current challenges. What is needed are criteria in terms of the original data of the given DAE. The projector based DAE analysis presented in this monograph is intended to address these questions.

We have been working on our theory of DAEs for quite some time. This theory has now achieved a certain maturity. Accordingly, it is time to record these developments in one coherent account. From the very beginning we were in the fortunate position to communicate with colleagues from all over the world, advancing different views on the topic, starting with Linda R. Petzold, Stephen L. Campbell, Werner C. Rheinboldt, Yuri E. Boyarintsev, Ernst Hairer, John C. Butcher and many others not mentioned here up to John D. Pryce, Ned Nedialkov, Andreas Griewank. We thank all of them for stimulating discussions.

For years, all of us have taught courses, held seminars, supervised diploma students and PhD students, and gained fruitful feedback, which has promoted the progress of our theory. We are indebted to all involved students and colleagues, most notably the PhD students.

Our work was inspired by several fascinating projects and long term cooperation, in particular with Roland England, Uwe Feldmann, Claus Führer, Michael Günther, Francesca Mazzia, Volker Mehrmann, Peter C. Müller, Peter Rentrop, Ewa Weinmüller, Renate Winkler.

We very much appreciate the joint work with Katalin Balla, who passed away too early in 2005, and the colleagues Michael Hanke, Immaculada Higueras, Galina Kurina, and Ricardo Riaza. All of them contributed essential ideas to the projector based DAE analysis.

We are indebted to the German Federal Ministry of Education and Research (BMBF) and the German Research Foundation (DFG), in particular the research center MATHEON in Berlin, for supporting our research in a lot of projects.

We would like to express our gratitude to many people for their support in the preparation of this volume. In particular we thank our colleague Jutta Kerger.

Last but not least, our special thanks are due to Achim Ilchmann and Timo Reis, the editors of the DAE Forum. We appreciate very much their competent counsel for improving the presentation of the theory.

We are under obligations to the staff of Springer for their careful assistance.

René Lamour Roswitha März Caren Tischendorf

Contents

Notations xv

Introduction xix

Part I Projector based approach

1 Linear constant coefficient DAEs 3
1.1 Regular DAEs and the Weierstraß–Kronecker form 3
1.2 Projector based decoupling of regular DAEs 10
1.2.1 Admissible matrix sequences and admissible projectors 10
1.2.2 Decoupling by admissible projectors 23
1.2.3 Complete decoupling 30
1.2.4 Hierarchy of projector sequences for constant matrix pencils 36
1.2.5 Compression to a generalized Weierstraß–Kronecker form . . 37
1.2.6 Admissible projectors for matrix pairs in a generalized Weierstraß–Kronecker form 41
1.3 Transformation invariance 45
1.4 Characterizing matrix pencils by admissible projectors 47
1.5 Properly stated leading term and solution space 50
1.6 Notes and references 52

2 Linear DAEs with variable coefficients 57
2.1 Properly stated leading terms 58
2.2 Admissible matrix function sequences 60
2.2.1 Basics 60
2.2.2 Admissible projector functions and characteristic values 65
2.2.3 Widely orthogonal projector functions 75
2.3 Invariants under transformations and refactorizations 79
2.4 Decoupling regular DAEs 86
2.4.1 Preliminary decoupling rearrangements 86

2.4.2 Regularity and basic decoupling of regular DAEs 90
2.4.3 Fine and complete decouplings 104
2.4.3.1 Index-1 case 104
2.4.3.2 Index-2 case 106
2.4.3.3 General benefits from fine decouplings 108
2.4.3.4 Existence of fine and complete decouplings 111
2.5 Hierarchy of admissible projector function sequences for linear DAEs 117
2.6 Fine regular DAEs 118
2.6.1 Fundamental solution matrices 119
2.6.2 Consistent initial values and flow structure 123
2.6.3 Stability issues 127
2.6.4 Characterizing admissible excitations and perturbation index 132
2.7 Specifications for regular standard form DAEs 137
2.8 The T-canonical form 140
2.9 Regularity intervals and critical points 146
2.10 Strangeness versus tractability 160
2.10.1 Canonical forms 160
2.10.2 Strangeness reduction 164
2.10.3 Projector based reduction 166
2.11 Generalized solutions 171
2.11.1 Measurable solutions 171
2.11.2 Distributional solutions 173
2.12 Notes and references 174

3 Nonlinear DAEs 183
3.1 Basic assumptions and notions 184
3.1.1 Properly involved derivative 184
3.1.2 Constraints and consistent initial values 187
3.1.3 Linearization 195
3.2 Admissible matrix function sequences and admissible projector functions 198
3.3 Regularity regions 208
3.4 Transformation invariance 224
3.5 Hessenberg form DAEs of arbitrary size 229
3.6 DAEs in circuit simulation 239
3.7 Local solvability 250
3.7.1 Index-1 DAEs 251
3.7.2 Index-2 DAEs 259
3.7.2.1 Advanced decoupling of linear index-2 DAEs 259
3.7.2.2 Nonlinear index-2 DAEs 261
3.7.2.3 Index reduction step 268
3.8 Advanced localization of regularity: including jet variables 272
3.9 Operator settings 281

3.9.1 Linear case . . . 283
3.9.2 Nonlinear case . . . 287
3.10 A glance at the standard approach via the derivative array and differentiation index . . . 290
3.11 Using structural peculiarities to ease models . . . 300
3.12 Regularity regions of DAEs with quasi-proper leading terms . . . 304
3.13 Notes and references . . . 307

Part II Index-1 DAEs: Analysis and numerical treatment

4 Analysis . . . 317
4.1 Basic assumptions and notions . . . 317
4.2 Structure and solvability of index-1 DAEs . . . 320
4.3 Consistent initial values . . . 334
4.4 Notes and references . . . 336

5 Numerical integration . . . 339
5.1 Basic idea . . . 340
5.2 Methods applied to ODEs and DAEs in standard form . . . 345
5.2.1 Backward differentiation formula . . . 345
5.2.2 Runge–Kutta method . . . 346
5.2.3 General linear method . . . 350
5.3 Methods applied to DAEs with a properly involved derivative . . . 352
5.3.1 Backward differentiation formula . . . 352
5.3.2 Runge–Kutta method . . . 353
5.3.3 General linear method . . . 355
5.4 When do decoupling and discretization commute? . . . 356
5.5 Convergence on compact intervals and error estimations . . . 361
5.5.1 Backward differentiation formula . . . 361
5.5.2 IRK(DAE) method . . . 364
5.5.3 General linear method . . . 369
5.6 Notes and references . . . 371

6 Stability issues . . . 375
6.1 Preliminaries concerning explicit ODEs . . . 375
6.2 Contractive DAEs and B-stable Runge–Kutta methods . . . 378
6.3 Dissipativity . . . 384
6.4 Lyapunov stability . . . 387
6.5 Notes and references . . . 394

Part III Computational aspects

7 Computational linear algebra aspects . . . 399
7.1 Image and nullspace projectors . . . 400

7.2 Matters of a properly stated leading term ... 402
7.3 The basic step of the sequence ... 404
7.3.1 Basis representation methods ... 406
7.3.2 Basis representation methods—Regular case ... 408
7.3.3 Projector representation method ... 409
7.4 Matrix function sequences ... 413
7.4.1 Stepping level by level ... 413
7.4.2 Involved version for the regular case ... 415
7.4.3 Computing characteristic values and index check ... 416

8 Aspects of the numerical treatment of higher index DAEs ... 419
8.1 Practical index calculation ... 419
8.2 Consistent initialization ... 424
8.3 Numerical integration ... 426
8.4 Notes and references ... 437

Part IV Advanced topics

9 Quasi-regular DAEs ... 441
9.1 Quasi-proper leading terms ... 441
9.2 Quasi-admissible matrix function sequences and admissible projector functions ... 446
9.3 Quasi-regularity ... 452
9.4 Linearization ... 455
9.5 A DAE transferable into SCF is quasi-regular ... 457
9.6 Decoupling of quasi-regular linear DAEs ... 462
9.7 Difficulties arising with the use of subnullspaces ... 468
9.8 Notes and references ... 471
9.9 Hierarchy of quasi-admissible projector function sequences for general nonlinear DAEs ... 475

10 Nonregular DAEs ... 477
10.1 The scope of interpretations ... 478
10.2 Linear DAEs ... 482
10.2.1 Tractability index ... 482
10.2.2 General decoupling ... 486
10.2.2.1 G_μ has full column rank ... 489
10.2.2.2 Tractability index 1, G_1 has a nontrivial nullspace ... 493
10.2.2.3 Tractability index 2, G_2 has a nontrivial nullspace ... 497
10.3 Underdetermined nonlinear DAEs ... 499
10.4 Notes and references ... 502

11 Minimization with constraints described by DAEs 505
11.1 Adjoint and self-adjoint DAEs 505
11.2 Extremal conditions and the optimality DAE 510
11.2.1 A necessary extremal condition and the optimality DAE 510
11.2.2 A particular sufficient extremal condition 520
11.3 Specification for controlled DAEs 522
11.4 Linear-quadratic optimal control and Riccati feedback solution 525
11.4.1 Sufficient and necessary extremal conditions 526
11.4.2 Riccati feedback solution 527
11.5 Notes and references 536

12 Abstract differential-algebraic equations 539
12.1 Index considerations for ADAEs 540
12.2 ADAE examples 544
12.2.1 Classical partial differential equations 545
12.2.1.1 Wave equation 545
12.2.1.2 Heat equation 546
12.2.2 A semi-explicit system with parabolic and elliptic parts 548
12.2.3 A coupled system of a PDE and Fredholm integral equations 552
12.2.4 A PDE and a DAE coupled by a restriction operator 554
12.3 Linear ADAEs with monotone operators 554
12.3.1 Basic functions and function spaces 555
12.3.2 Galerkin approach 558
12.3.3 Solvability 564
12.3.4 Continuous dependence on the data 571
12.3.5 Strong convergence of the Galerkin method 574
12.4 Notes and references 577

Appendices 581

A Linear algebra – basics 581
A.1 Projectors and subspaces 581
A.2 Generalized inverses 589
A.3 Parameter-dependent matrices and projectors 591
A.4 Variable subspaces 594

B Technical computations 599
B.1 Proof of Lemma 2.12 599
B.2 Proof of Lemma 2.41 605
B.3 Admissible projectors for $Nx' + x = r$ 612

C Analysis 627
C.1 A representation result 627

C.2 ODEs 629
C.3 Basics for evolution equations 632

References 637

Index 647

Notations

Abbreviations

ADAE	abstract DAE
BDF	backward differentiation formula
DAE	differential-algebraic equation
GLM	general linear method
IERODE	inherent explicit regular ODE
IESODE	inherent explicit singular ODE
IVP	initial value problem
MNA	modified nodal analysis
ODE	ordinary differential equation
PDAE	partial DAE
SCF	standard canonical form
SSCF	strong SCF

Common notation

$\mathbb{N}$	natural numbers
$\mathbb{R}$	real numbers
$\mathbb{C}$	complex numbers
$\mathbb{K}$	alternatively $\mathbb{R}$ or $\mathbb{C}$
$\mathbb{K}^n$	n-dimensional vector space
$M \in \mathbb{K}^{m,n}$	matrix with m rows and n columns
$M \in L(\mathbb{K}^n, \mathbb{K}^m)$	linear mapping from $\mathbb{K}^n$ into $\mathbb{K}^m$, also for $M \in \mathbb{K}^{m,n}$
$L(\mathbb{K}^m)$	shorthand for $L(\mathbb{K}^m, \mathbb{K}^m)$
M^{T}	transposed matrix
M^*	transposed matrix with real or complex conjugate entries
M^{-1}	inverse matrix
M^-	reflexive generalized inverse of M
M^+	Moore–Penrose inverse of M
$\ker M$	kernel of M, $\ker M = \{z \mid Mz = 0\}$

$\operatorname{im} M$	image of M, $\operatorname{im} M = \{z \mid z = My,\ y \in \mathbb{R}^n\}$
$\operatorname{ind} M$	index of M, $\operatorname{ind} M = \min\{k : \ker M^k = \ker M^{k+1}\}$
$\operatorname{rank} M$	rank of M
$\det M$	determinant of M
span	linear hull of a set of vectors
dim	dimension of a (sub)space
diag	diagonal matrix
$\{0\}$	set containing the zero element only
$M \cdot \mathcal{N}$	$= \{z \mid z = My,\ y \in \mathcal{N}\}$
$\forall$	for all
$\perp$	orthogonal set, $\mathcal{N}^\perp = \{z \mid \langle n, z\rangle = 0, \forall n \in \mathcal{N}\}$
$\otimes$	Kronecker product
$\oplus$	direct sum
$\ominus$	$\mathcal{X} = \mathcal{N}_i \ominus \mathcal{N}_j \Leftrightarrow \mathcal{N}_i = \mathcal{X} \oplus \mathcal{N}_j$
$\{A, B\}$	ordered pair
$\lvert \cdot \rvert$	vector and matrix norms in $\mathbb{R}^m$
$\lvert\lvert \cdot \rvert\rvert$	function norm
$\langle \cdot, \cdot \rangle$	scalar product in $\mathbb{K}^m$, dual pairing
$(\cdot \mid \cdot)_H$	scalar product in Hilbert space H
I, I_d	identity matrix (of dimension d)
$\mathcal{I}$	interval of independent variable
$(\ )'$	total time derivative, total derivative in jet variables
$(\ )_x$	(partial) derivative with respect to x
$\mathcal{C}(\mathcal{I}, \mathbb{R}^m)$	set of continuous functions
$\mathcal{C}^k(\mathcal{I}, \mathbb{R}^m)$	set of k-times continuously differentiable functions
$L_2(\mathcal{I}, \mathbb{R}^m)$	Lebesgue space
$H^1(\mathcal{I}, \mathbb{R}^m)$	Sobolev space

Special notation

$\mathcal{M}_0(t)$	obvious constraint
G_i	member of admissible matrix function sequence
r_i	$\operatorname{rank} G_i$, see Definition 1.17
S_j	$S_j = \ker \mathcal{W}_j B$, see Theorem 2.8 and following pages
N_j	$N_j = \ker G_j$, in Chapter 9: N_j subspace of $\ker G_j$
$\widehat{N}_i$	intersection: $N_0 + \cdots + N_{i-1} \cap N_i$, see (1.12)
N_{can}	canonical subspace, see Definition 2.36
$N_{can\,\mu}$	canonical subspace of an index μ DAE
S_{can}	canonical subspace (Definition 2.36)
$M_{can,q}$	set of consistent values, see (2.98)
$\mathcal{I}_{reg}$	set of regular points, see Definition 2.74
$X(\cdot, t_0)$	fundamental solution matrix normalized at t_0
dom_f	definition domain of f
$\mathcal{C}^k$-subspace	smooth subspace (cf. Section A.4)
$\mathcal{C}^\nu_*(\mathcal{G})$	set of reference functions, see Definition 3.17

$\mathcal{C}_D^1$	$\mathcal{C}_D^1(\mathcal{I},\mathbb{R}^m) := \{x \in \mathcal{C}(\mathcal{I},\mathbb{R}^m) : Dx \in \mathcal{C}^1(\mathcal{I},\mathbb{R}^n)\}$, see (1.78)
H_D^1	$H_D^1(\mathcal{I},\mathbb{R}^m) := \{x \in \mathcal{L}_2(\mathcal{I},\mathbb{R}^m) :\ Dx \in H^1(\mathcal{I},\mathbb{R}^n)\}$
$\mathcal{C}^{ind\ \mu}$	function space, see (2.104)
$\mathcal{G}$	regularity region

For projectors we usually apply the following notation:

Q	nullspace projector of a matrix G, $\operatorname{im} Q = \ker G$, $GQ = 0$
P	complementary projector, $P = I - Q$, $GP = G$
$\mathcal{W}$	projector along the image of G, $\ker \mathcal{W} = \operatorname{im} G$, $\mathcal{W}G = 0$
$P_i \cdots P_j$	ordered product, $\prod_{k=i}^{j} P_k$
Π_i	$\Pi_i := P_0 P_1 \cdots P_i$
Π_{can} $\Pi_{can\ \mu}$	canonical projector (of an index-μ DAE)
P_{dich}	dichotomic projector, see Definition 2.56

Introduction

Ordinary differential equations (ODEs) define relations concerning function values and derivative values of an unknown vector valued function in one real independent variable often called time and denoted by t. An explicit ODE

$$x'(t) = g(x(t),t)$$

displays the derivative value $x'(t)$ explicitly in terms of t and $x(t)$. An implicit ODE

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

is said to be regular, if all its line-elements (x^1,x,t) are regular. A triple (x^1,x,t) belonging to the domain of interest is said to be a regular line-element of the ODE, if $f_{x^1}(x^1,x,t)$ is a nonsingular matrix, and otherwise a singular line-element. This means, in the case of a regular ODE, the derivative value $x'(t)$ is again fully determined in terms of t and $x(t)$, but in an implicit manner.

An ODE having a singular line-element is said to be a singular ODE. In turn, singular ODEs comprise quite different classes of equations. For instance, the linear ODE

$$tx'(t) - Mx(t) = 0$$

accommodates both regular line-elements for $t \neq 0$ and singular ones for $t = 0$. In contrast, the linear ODE

$$\begin{bmatrix} 1&0&0&0&0 \\ 0&0&1&0&0 \\ 0&0&0&1&0 \\ 0&0&0&0&1 \\ 0&0&0&0&0 \end{bmatrix} x'(t) + \begin{bmatrix} -\alpha & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} x(t) - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \gamma(t) \end{bmatrix} = 0 \tag{0.1}$$

has solely singular line-elements. A closer look at the solution flow of the last two ODEs shows a considerable disparity.

The ODE (0.1) serves as a prototype of a differential-algebraic equation (DAE). The related equation $f(x^1,x,t)=0$ determines the components x^1_1, x^1_3, x^1_4, and x^1_5 of x^1 in terms of x and t. The component x^1_2 is not at all given. In addition, there arises the consistency condition $x_5-\gamma(t)=0$ which restricts the flow.

DAEs constitute—in whatever form they are given—somehow uniformly singular ODEs: In common with all ODEs, they define relations concerning function values and derivative values of an unknown vector valued function in one real independent variable. However, in contrast to explicit ODEs, in DAEs these relations are implicit, and, in contrast to regular implicit ODEs, these relations determine just a part of the derivative values. A DAE is an implicit ODE which has solely singular line-elements.

The solutions of the special DAE (0.1) feature an ambivalent nature. On the one hand they are close to solutions of regular ODEs in the sense that they depend smoothly on consistent initial data. On the other hand, tiny changes of γ may yield monstrous variations of the solutions, and the solution varies discontinuously with respect to those changes. We refer to the figures in Example 1.5 to gain an impression of this ill-posed behavior.

The ambivalent nature of their solutions distinguishes DAE as being extraordinary to a certain extent.

DAEs began to attract significant research interest in applied and numerical mathematics in the early 1980s, no more than about three decades ago. In this relatively short time, DAEs have become a widely acknowledged tool to model processes subject to constraints, in order to simulate and to control these processes in various application fields.

The two traditional physical application areas, network simulation in electronics and the simulation of multibody mechanics, are repeatedly addressed in textbooks and surveys (e.g. [96, 25, 189]). Special monographs [194, 63, 188] and much work in numerical analysis are devoted to these particular problems. These two application areas and related fields in science and engineering can also be seen as the most important impetus to begin with systematic DAE research, since difficulties and failures in respective numerical simulations have provoked the analysis of these equations first.

The equations describing electrical networks have the form

$$A(d(x(t),t))'+b(x(t),t)=0, \tag{0.2}$$

with a singular constant matrix A, whereas constrained multibody dynamics is described by equations showing the particular structure

$$x_1'(t)+b_1(x_1(t),x_2(t),x_3(t),t)=0, \tag{0.3}$$
$$x_2'(t)+b_2(x_1(t),x_2(t),t)=0, \tag{0.4}$$
$$b_3(x_2(t),t)=0. \tag{0.5}$$

Those DAEs usually have large dimension. Multibody systems often comprise hundreds of equations and electric network systems even gather up to several millions of equations.

Many further physical systems are naturally described as DAEs, for instance, chemical process modeling, [209]. We agree with [189, p. 192] that DAEs arise probably more often than (regular) ODEs, and many of the well-known ODEs in application are actually DAEs that have been additionally explicitly reduced to ODE form.

Further DAEs arise in mathematics, in particular, as intermediate reduced models in singular perturbation theory, as extremal conditions in optimization and control, and by means of semidiscretization of partial differential equation systems.

Besides the traditional application fields, conducted by the generally increasing role of numerical simulation in science and technology, currently more and more new applications come along, in which different physical components are coupled via a network.

We believe that DAEs and their more abstract versions in infinite-dimensional spaces comprise *great potential for future mathematical modeling*. To an increasingly large extent, in applications, DAEs are automatically generated, often by coupling various subsystems, with large dimensions, but *without manifested mathematically useful structures*. Different modeling approaches may result in different kinds of DAEs. Automatic generation and coupling of various tools may yield quite opaque DAEs. Altogether, this produces the challenging task to *bring to light and to characterize the inherent mathematical structure of DAEs*, to provide test criteria such as index observers and eventually hints for creating better qualified model modifications. For a reliable practical treatment, which is the eventual aim, for numerical simulation, sensitivity analysis, optimization and control, and last but not least practical upgrading models, one needs pertinent information concerning the mathematical structure. Otherwise their procedures may fail or, so much the worse, generate wrong results. In consequence, providing practical assessment tools to uncover and to monitor mathematical DAE structures is one of the actual challenges. What are needed are criteria in terms of the original data of the given DAE. The projector based DAE analysis presented in this monograph is intended to address these questions.

Though DAEs have been popular among numerical analysts and in various application fields, so far they play only a marginal role in contiguous fields such as nonlinear analysis and dynamical systems. However, an input from those fields would be desirable. It seems, responsible for this shortage is the quite common view of DAEs as in essence nothing other than implicitly written regular ODEs or vector fields on manifolds, making some difficulties merely in numerical integration. The latter somehow biased opinion is still going strong. It is fortified by the fact that almost all approaches to DAEs suppose that the DAE is eventually reducible to an ODE as a basic principle. This opinion is summarized in [189, p. 191] as follows: *It is a fact, not a mere point of view, that a DAE eventually reduces to an ODE on a manifold. The attitude of acknowledging this fact from the outset leads to a reduc-*

tion procedure suitable for the investigation of many problems .... The mechanism of the geometric reduction procedure completely elucidates the "algebraic" and the "differential" aspects of a DAE. The algebraic part consists in the characterization of the manifold over which the DAE becomes an ODE and, of course, the differential part provides the reduced ODE. Also in [130] the explicit reduction of the general DAE

$$\mathfrak{f}(x'(t), x(t), t) = 0, \tag{0.6}$$

with a singular partial Jacobian $\mathfrak{f}_{x'}$, into a special reduced form plays a central role. Both monographs [189, 130] concentrate on related reduction procedures which naturally suppose higher partial derivatives of the function $\mathfrak{f}$, either to provide sequences of smooth (sub)manifolds or to utilize a so-called derivative array system. The differential geometric approach and the reduction procedures represent powerful tools to analyze and to solve DAEs. Having said that, we wonder about the misleading character of this purely geometric view, which underlines the closedness to regular ODEs, but loses sight of the ill-posed feature.

So far, most research concerning general DAEs is addressed to equation (0.6), and hence we call this equation a *DAE in standard form*. Usually, a solution is then supposed to be at least continuously differentiable.
In contrast, in the present monograph we investigate equations of the form

$$f((d(x(t), t))', x(t), t) = 0, \tag{0.7}$$

which show the derivative term involved by means of an extra function d. We see the network equation (0.2) as the antetype of this form. Also the system (0.3)–(0.5) has this form

$$\begin{bmatrix} I & 0 \\ 0 & I \\ 0 & 0 \end{bmatrix} \left(\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \right)' + \begin{bmatrix} b_1(x_1(t), x_2(t), x_3(t), t) \\ b_2(x_1(t), x_2(t), t) \\ b_3(x_2(t), t) \end{bmatrix} = 0 \tag{0.8}$$

a priori. It appears that in applications actually DAEs in the form (0.7) arise, which precisely indicates the involved derivatives. The DAE form (0.7) is comfortable; it involves the derivative by the extra nonlinear function d, whereby $x(t) \in \mathbb{R}^m$ and $d(x(t), t) \in \mathbb{R}^n$ may have different sizes, as is the case in (0.8). A particular instance of DAEs (0.7) is given by the so-called *conservative form* DAEs [52]. Once again, the idea for version (0.7) originates from circuit simulation problems, in which this form is well approved (e.g. [75, 168]).

However, though equation (0.7) represents a more precise model, one often transforms it to standard form (0.6), which allows to apply results and tools from differential geometry, numerical ODE methods, and ODE software.

Turning from the model (0.7) to a standard form DAE one veils the explicit precise information concerning the derivative part. With this background, we are confronted with the question of what a DAE solution should be. Following the classical sense of differential equations, we ask for continuous functions being *as smooth as necessary*, which satisfy the DAE pointwise on the interval of interest. This is a common understanding. However, there are different opinions on the meaning of the appropriate smoothness. Having regular ODEs in mind one considers continuously differentiable functions $x(\cdot)$ to be the right candidates for solutions. Up to now, most DAE researchers adopt this understanding of the solution which is supported by the standard DAE formulation. Furthermore, intending to apply formal integrability concepts, differential geometry and derivative array approaches one is led to yet another higher smoothness requirement. In contrast, the multibody system (0.8) suggests, as solutions, continuous functions $x(\cdot)$ having just continuously differentiable components $x_1(\cdot)$ and $x_2(\cdot)$.

An extra matrix figuring out the derivative term was already used much earlier (e.g. [153, 152, 154]); however, this approach did not win much recognition at that time. Instead, the following interpretation of standard form DAEs (e.g. [96]) has been accepted to a larger extent: Assuming the nullspace of the partial Jacobian $\mathfrak{f}_{x'}(x',x,t)$ associated with the standard form DAE (0.6) to be a $\mathcal{C}^1$-subspace, and to be independent of the variables x' and x, one interprets the standard form DAE (0.6) as a short description of the equation

$$\mathfrak{f}((P(t)x(t))' - P'(t)x(t), x(t), t) = 0, \tag{0.9}$$

whereby $P(\cdot)$ denotes any continuously differentiable projector valued function such that the nullspaces $\ker P(\cdot)$ and $\ker \mathfrak{f}_{x'}(x',x,\cdot)$ coincide. This approach is aligned with continuous solutions $x(\cdot)$ having just continuously differentiable products $(Px)(\cdot)$. Most applications yield even constant nullspaces $\ker \mathfrak{f}_{x'}$, and hence constant projector functions P as well. In particular, this is the case for the network equations (0.2) and the multibody systems (0.8).

In general, for a DAE given in the form (0.7), a solution $x(\cdot)$ should be a continuous function such that the superposition $u(\cdot) := d(x(\cdot),\cdot)$ is continuously differentiable. For the particular system (0.8) this means that the components $x_1(\cdot)$ and $x_2(\cdot)$ are continuously differentiable, whereas one accepts a continuous $x_3(\cdot)$.

The question in which way the data functions f and d should be related to each other leads to the notions of *DAEs with properly stated leading term or properly involved derivative*, but also to *DAEs with quasi-proper leading term*. During the last 15 years, the idea of using an extra function housing the derivative part within a DAE has been emphatically pursued. This discussion amounts to the content of this monograph. Formulating DAEs with properly stated leading term yields, in particular, symmetries of linear DAEs and their adjoints, and further favorable consequences concerning optimization problems with DAE constraints. Not surprisingly, numerical discretization methods may perform better than for standard form DAEs. And last, but not least, this approach allows for appropriate generalizations to apply to abstract differential- algebraic systems in Hilbert spaces enclosing PDAEs. We think that, right from the design or modeling stage, it makes sense to look for properly involved derivatives.

This monograph comprises an elaborate analysis of DAEs (0.7), which is accompanied by the consideration of essential numerical aspects. We regard DAEs from an analytical point of view, rather than from a geometric one. Our main objective consists in the structural and qualitative characterization of DAEs as they are given a priori, without supposing any knowledge concerning solutions and constraints. Afterwards, having the required knowledge of the DAE structure, also solvability assertions follow. Only then do we access the constraints. In contrast, other approaches concede full priority of providing constraints and solutions, as well as transformations into a special form, which amounts to solving the DAE.

We believe in the great potential of our concept in view of the further analysis of classical DAEs and their extensions to abstract DAEs in function spaces. We do not at all apply derivative arrays and prolongated systems, which are commonly used in DAE theory. Instead, certain admissible matrix function sequences and smartly chosen admissible projector functions formed only from the first partial derivatives of the given data function play their role as basic tools. Thereby, continuity properties of projector functions depending on several variables play their role, which is not given if one works instead with basises. All in all, this allows an analysis on a low smoothness level. We pursue a fundamentally alternative approach and present the first rigorous structural analysis of general DAEs in their originally given form without the use of derivative arrays, without supposing any knowledge concerning constraints and solutions.

The concept of a projector based analysis of general DAEs was sketched first in [160, 171, 48], but it has taken its time to mature. Now we come up with a unique general theory capturing constant coefficient linear problems, variable coefficient linear problems and fully nonlinear problems in a hierarchic way. We address a further generalization to abstract DAEs. It seems, after having climbed the (at times seemingly pathless) mountain of projectors, we are given transparency and beautiful convenience. By now the projector based analysis is approved to be a prospective way to investigate DAEs and also to yield reasonable open questions for future research.

The central idea of the present monograph consists in a rigorous definition of regularity of a DAE, accompanied with certain characteristic values including the tractability index, which is related to an open subset of the definition domain of the data function f, a so-called *regularity region*. Regularity is shown to be stable with respect to perturbations. Close relations of regularity regions and linearizations are proved. In general, one has to expect that the definition domain of f decomposes into several regularity regions whose borders consist of critical points. Solutions do not necessarily stay in one of these regions; solutions may cross the borders and undergo bifurcation, etc.

The larger part of the presented material is new and as yet unpublished. Parts were earlier published in journals, and just the regular linear DAE framework (also critical points in this context) is available in the book [194].

The following basic types of DAEs can reasonably be discerned:

✓ fully implicit nonlinear DAE with nonlinear derivative term

$$f((d(x(t),t))',x(t),t) = 0, \tag{0.10}$$

✓ fully implicit nonlinear DAE with linear derivative term

$$f((D(t)x(t))', x(t), t) = 0, \tag{0.11}$$

✓ quasi-linear DAE with nonlinear derivative term (involved linearly)

$$A(x(t),t)(d(x(t),t))' + b(x(t),t) = 0, \tag{0.12}$$

✓ quasi-linear DAE with linear derivative term

$$A(x(t),t)(D(t)x(t))' + b(x(t),t) = 0, \tag{0.13}$$

✓ linear DAE with variable coefficients

$$A(t)(D(t)x(t))' + B(t)x(t) = q(t), \tag{0.14}$$

✓ linear DAE with constant coefficients

$$A(Dx(t))' + Bx(t) = q(t), \tag{0.15}$$

✓ semi-implicit DAE with explicitly given derivative-free equation

$$f_1((d(x(t),t))', x(t), t) = 0, \tag{0.16}$$
$$f_2(x(t),t) = 0, \tag{0.17}$$

✓ semi-implicit DAE with explicitly partitioned variable and explicitly given derivative-free equation

$$f_1(x_1'(t), x_1(t), x_2(t), t) = 0, \tag{0.18}$$
$$f_2(x_1(t), x_2(t), t) = 0, \tag{0.19}$$

✓ semi-explicit DAE with explicitly partitioned variable and explicitly given derivative-free equation

$$x_1'(t) + b_1(x_1(t), x_2(t), t) = 0, \tag{0.20}$$
$$b_2(x_1(t), x_2(t), t) = 0. \tag{0.21}$$

So-called Hessenberg form DAEs of size r, which are described in Section 3.5, form further subclasses of semi-explicit DAEs. For instance, the DAE (0.8) has Hessenberg form of size 3. Note that much work developed to treat higher index DAEs is actually limited to Hessenberg form DAEs of size 2 or 3.

The presentation is divided into Part I to Part IV followed by Appendices A, B, and C.

Part I describes the core of the projector based DAE analysis: the construction of admissible matrix function sequences associated by admissible projector functions and the notion of regularity regions.

Chapter 1 deals with constant coefficient DAEs and matrix pencils only. We reconsider algebraic features and introduce into the projector framework. This can be skipped by readers familiar with the basic linear algebra including projectors.

The more extensive Chapter 2 provides the reader with admissible matrix function sequences and the resulting constructive projector based decouplings. With this background, a comprehensive linear theory is developed, including qualitative flow characterizations of regular DAEs, the rigorous description of admissible excitations, and also relations to several canonical forms and the strangeness index.

Chapter 3 contains the main constructions and assertions concerning general regular nonlinear DAEs, in particular the regularity regions and the practically important theorem concerning linearizations. It is recommended to take a look to Chapter 2 before reading Chapter 3.

We emphasize the hierarchical organization of Part I. The admissible matrix function sequences built for the nonlinear DAE (0.10) generalize those for the linear DAE (0.14) with variable coefficients, which, in turn, represent a generalization of the matrix sequences made for constant coefficient DAEs (0.15).

Part IV continues the hierarchy in view of different further aspects. Chapter 9 about quasi-regular DAEs (0.10) incorporates a generalization which relaxes the constant-rank conditions supporting admissible matrix function sequences. Chapter 10 on nonregular DAEs (0.11) allows a different number of equations and of unknown components. Finally, in Chapter 12, we describe abstract DAEs in infinite-dimensional spaces and include PDAEs.

Part IV contains the additional Chapter 11 conveying results on minimization with DAE constraints obtained by means of the projector based technique.

Part II is a self-contained index-1 script. It comprises in its three chapters the analysis of regular index-1 DAEs (0.11) and their numerical integration, addressing also stability topics such as contractivity and stability in Lyapunov's sense. Part II constitutes in essence an up-to-date improved and completed version of the early book [96]. While the latter is devoted to standard form DAEs via the interpretation (0.9), now the more general equations (0.11) are addressed.

Part III adheres to Part I giving an elaborate account of computational methods concerning the practical construction of projectors and that of admissible projector functions in Chapter 7. A second chapter discusses several aspects of the numerical treatment of regular higher index DAEs such as consistent initialization and numerical integration.

Appendix B contains technically involved costly proofs. Appendices A and C collect and provide basic material concerning linear algebra and analysis, for instance the frequently used $\mathcal{C}^1$-subspaces.

Plenty of reproducible small academic examples are integrated into the explanations for easier reading, illustrating and confirming the features under consideration. To this end, we emphasize that those examples are always too simple. They bring to light special features, but they do not really reflect the complexity of DAEs.

The material of this monograph is much too comprehensive to be taught in a standard graduate course. However different combinations of selected chapters should be well suited for those courses. In particular, we recommend the following:

- Projector based DAE analysis (Part I, possibly without Chapter 1).
- Analysis of index-1 DAEs and their numerical treatment (Part II, possibly plus Chapter 8).
- Matrix pencils, theoretical and practical decouplings (Chapters 1 and 7).
- General linear DAEs (Chapter 2, material on the linear DAEs of Chapters 10 and 9).

Advanced courses communicating Chapter 12 or Chapter 11 could be given to students well grounded in DAE basics (Parts I and II) and partial differential equations, respectively optimization.

Part I
Projector based approach

Part I describes the core of the projector based DAE analysis, the construction of admissible matrix function sequences and the notions of regular points and regularity regions of general DAEs

$$f((d(x(t),t))',x(t),t)=0$$

in a hierarchical manner starting with constant coefficient linear DAEs, then turning to linear DAEs with variable coefficients, and, finally, considering fully implicit DAEs.

Chapter 1 deals with constant coefficient DAEs and matrix pencils. We reconsider algebraic features and introduce them into the projector framework. This shows how the structure of the Weierstraß–Kronecker form of a regular matrix pencil can be depicted by means of admissible projectors.

The extensive Chapter 2 on linear DAEs with variable coefficients characterizes regular DAEs by means of admissible matrix function sequences and associated projectors and provides constructive projector based decouplings of regular linear DAEs.

Then, with this background, a comprehensive linear theory of regular DAEs is developed, including qualitative flow properties and a rigorous description of admissible excitations. Moreover, relations to several canonical forms and other index notions are addressed.

Chapter 3 contains the main constructions and assertions concerning general regular nonlinear DAEs, in particular the regularity regions and the practically important theorem concerning linearizations. Also local solvability assertions and perturbation results are proved.

We emphasize the hierarchical organization of the approach. The admissible matrix function sequences built for the nonlinear DAE (0.10) generalize those for the linear DAE (0.14) with variable coefficients, which, in turn, represent a generalization of the matrix sequences made for constant coefficient DAEs (0.15). Part IV continues the hierarchy with respect to different views.

Chapter 1
Linear constant coefficient DAEs

Linear DAEs with constant coefficients have been well understood by way of the theory of matrix pencils for quite a long time, and this is the reason why they are only briefly addressed in monographs. We consider them in detail here, not because we believe that the related linear algebra has to be invented anew, but as we intend to give a sort of guide for the subsequent extensive discussion of linear DAEs with time-varying coefficients and of nonlinear DAEs.

This chapter is organized as follows. Section 1.1 records well-known facts on regular matrix pairs and describes the structure of the related DAEs. The other sections serve as an introduction to the projector based analysis. Section 1.2 first provides the basic material of this analysis: the admissible matrix sequences and the accompanying admissible projectors and characteristic values in Subsection 1.2.1, the decoupling of regular DAEs by arbitrary admissible projectors in Subsection 1.2.2, and the complete decoupling in Subsection 1.2.3. The two subsequent Subsections 1.2.5 and 1.2.6 are to clarify the relations to the Weierstraß–Kronecker form. Section 1.3 provides the main result concerning the high consistency of the projector based approach and the DAE structure by the Weierstraß–Kronecker form, while Section 1.4 collects practically useful details on the topic. Section 1.5 develops proper formulations of the leading term of the DAE by means of two well-matched matrices. The chapter ends with notes and references.

1.1 Regular DAEs and the Weierstraß–Kronecker form

In this section we deal with the equation

$$Ex'(t) + Fx(t) = q(t),\ t \in \mathcal{I}, \tag{1.1}$$

formed by the ordered pair $\{E, F\}$ of real valued $m \times m$ matrices E, F. For given functions $q : \mathcal{I} \to \mathbb{R}^m$ being at least continuous on the interval $\mathcal{I} \subseteq \mathbb{R}$, we are looking for continuous solutions $x : \mathcal{I} \to \mathbb{R}^m$ having a continuously differentiable com-

R. Lamour et al., *Differential-Algebraic Equations: A Projector Based Analysis*,
Differential-Algebraic Equations Forum, DOI 10.1007/978-3-642-27555-5_1,

ponent Ex. We use the notation $Ex'(t)$ for $(Ex)'(t)$. Special interest is directed to homogeneous equations

$$Ex'(t)+Fx(t)=0,\ \ t\in\mathbb{R}. \tag{1.2}$$

For $E=I$, the special case of explicit ODEs is covered. Now, in the more general setting, the ansatz $x_*(t)=e^{\lambda_* t}z_*$ well-known for explicit ODEs, yields

$$Ex_*'(t)+Fx_*(t)=e^{\lambda_* t}(\lambda_* E+F)z_*.$$

Hence, x_* is a nontrivial particular solution of the DAE (1.2) if λ_* is a zero of the polynomial $p(\lambda):=\det(\lambda E+F)$, and $z_*\neq 0$ satisfies the relation $(\lambda_* E+F)z_*=0$. Then λ_* and z_* are called generalized eigenvalue and eigenvector, respectively. This shows the meaning of the polynomial $p(\lambda)$ and the related family of matrices $\lambda E+F$ named the *matrix pencil* formed by $\{E,F\}$.

Example 1.1 (A solvable DAE). The DAE

$$\begin{aligned} x_1'-x_1&=0,\\ x_2'+x_3&=0,\\ x_2&=0, \end{aligned}$$

is given by the matrices

$$E=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix} \text{ and } F=\begin{bmatrix}-1&0&0\\0&0&1\\0&1&0\end{bmatrix},$$

yielding

$$p(\lambda)=\det(\lambda E+F)=\det\begin{bmatrix}\lambda-1&0&0\\0&\lambda&1\\0&1&0\end{bmatrix}=1-\lambda.$$

The value $\lambda_*=1$ is a generalized eigenvalue and the vector $z_*=(1\,0\,0)^{\mathrm{T}}$ is a generalized eigenvector. Obviously, $x_*(t)=e^{\lambda_* t}z_*=(e^t\,0\,0)^{\mathrm{T}}$ is a nontrivial solution of the differential-algebraic equation. □

If E is nonsingular, the homogeneous equation (1.2) represents an implicit regular ODE and its fundamental solution system forms an m-dimensional subspace in $\mathcal{C}^1(\mathcal{I},\mathbb{R}^m)$. What happens if E is singular? Is there a class of equations, such that equation (1.2) has a finite-dimensional solution space? The answer is closely related to the notion of regularity.

Definition 1.2. Given any ordered pair $\{E,F\}$ of matrices $E,F\in L(\mathbb{R}^m)$, the matrix pencil $\lambda E+F$ is said to be *regular* if the polynomial $p(\lambda):=\det(\lambda E+F)$ does not vanish identically. Otherwise the matrix pencil is said to be *singular*.
Both the ordered pair $\{E,F\}$ and the DAE (1.1) are said to be *regular* if the accompanying matrix pencil is regular, and otherwise *nonregular*.

A pair $\{E,F\}$ with a nonsingular matrix E is always regular, and its polynomial $p(\lambda)$ is of degree m. In the case of a singular matrix E, the polynomial degree is lower as demonstrated in Example 1.1.

Proposition 1.3. *For any regular pair $\{E,F\}$, $E,F \in L(\mathbb{R}^m)$, there exist nonsingular matrices $L,K \in L(\mathbb{R}^m)$ and integers $0 \le l \le m$, $0 \le \mu \le l$, such that*

$$LEK = \begin{bmatrix} I & \\ & N \end{bmatrix} \begin{matrix} \}m-l \\ \}l \end{matrix} \quad , \quad LFK = \begin{bmatrix} W & \\ & I \end{bmatrix} \begin{matrix} \}m-l \\ \}l \end{matrix} \quad . \tag{1.3}$$

Thereby, N is absent if $l=0$, and otherwise N is nilpotent of order μ, i.e., $N^\mu = 0$, $N^{\mu-1} \neq 0$. The integers l and μ as well as the eigenstructure of the blocks N and W are uniquely determined by the pair $\{E,F\}$.

Proof. If E is nonsingular, we simply put $l=0$, $L=E^{-1}$, $K=I$ and the assertion is true.
Assume E to be singular. Since $\{E,F\}$ is a regular pair, there is a number $c \in \mathbb{R}$ such that $cE+F$ is nonsingular. Form $\tilde{E} := (cE+F)^{-1}E$, $\tilde{F} := (cE+F)^{-1}F = I - c\tilde{E}$, $\mu = \operatorname{ind}\tilde{E}$, $r = \operatorname{rank}\tilde{E}^\mu$, $S = [s_1 \ldots s_m]$ with $s_1,\ldots,s_r$ and $s_{r+1},\ldots,s_m$ being bases of $\operatorname{im}\tilde{E}^\mu$ and $\ker\tilde{E}^\mu$, respectively. Lemma A.11 provides the special structure of the product $S^{-1}\tilde{E}S$, namely,

$$S^{-1}\tilde{E}S = \begin{bmatrix} \tilde{M} & 0 \\ 0 & \tilde{N} \end{bmatrix},$$

with a nonsingular $r \times r$ block $\tilde{M}$ and a nilpotent $(m-r)\times(m-r)$ block $\tilde{N}$. $\tilde{N}$ has nilpotency index μ. Compute

$$S^{-1}\tilde{F}S = I - cS^{-1}\tilde{E}S = \begin{bmatrix} I-c\tilde{M} & 0 \\ 0 & I-c\tilde{N} \end{bmatrix}.$$

The block $I - c\tilde{N}$ is nonsingular due to the nilpotency of $\tilde{N}$. Denote

$$L := \begin{bmatrix} \tilde{M}^{-1} & 0 \\ 0 & (I-c\tilde{N})^{-1} \end{bmatrix} S^{-1}(cE+F)^{-1},$$
$$K := S, \qquad N := (I-c\tilde{N})^{-1}\tilde{N}, \qquad W := \tilde{M}^{-1} - cI,$$

so that we arrive at the representation

$$LEK = \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}, \quad LFK = \begin{bmatrix} W & 0 \\ 0 & I \end{bmatrix}.$$

Since $\tilde{N}$ and $(I-c\tilde{N})^{-1}$ commute, one has

$$N^l = ((I-c\tilde{N})^{-1}\tilde{N})^l = ((I-c\tilde{N})^{-1})^l\tilde{N}^l,$$

and N inherits the nilpotency of $\tilde{N}$. Thus, $N^\mu = 0$ and $N^{\mu-1} \neq 0$. Put $l := m-r$. It remains to verify that the integers l and μ as well as the eigenstructure of N and

W are independent of the transformations L and K. Assume that there is a further collection $\tilde{l}, \tilde{\mu}, \tilde{L}, \tilde{K}$, $\tilde{r} = m - \tilde{l}$ such that

$$\tilde{L}E\tilde{K} = \begin{bmatrix} I_{\tilde{r}} & 0 \\ 0 & \tilde{N} \end{bmatrix}, \quad \tilde{L}F\tilde{K} = \begin{bmatrix} \tilde{W} & 0 \\ 0 & I_{\tilde{l}} \end{bmatrix}.$$

Considering the degree of the polynomial

$$\begin{aligned} p(\lambda) = \det(\lambda E + F) &= \det(L^{-1}) \det(\lambda I_r + W) \det(K^{-1}) \\ &= \det(\tilde{L}^{-1}) \det(\lambda I_{\tilde{r}} + \tilde{W}) \det(\tilde{K}^{-1}) \end{aligned}$$

we realize that the values r and $\tilde{r}$ must coincide, hence $l = \tilde{l}$. Introducing $U := \tilde{L}L^{-1}$ and $V := \tilde{K}^{-1}K$ one has

$$U \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} = \tilde{L}EK = \begin{bmatrix} I & 0 \\ 0 & \tilde{N} \end{bmatrix} V, \quad U \begin{bmatrix} W & 0 \\ 0 & I \end{bmatrix} = \tilde{L}FK = \begin{bmatrix} \tilde{W} & 0 \\ 0 & I \end{bmatrix} V,$$

and, in detail,

$$\begin{bmatrix} U_{11} & U_{12}N \\ U_{21} & U_{22}N \end{bmatrix} = \begin{bmatrix} V_{11} & V_{12} \\ \tilde{N}V_{21} & \tilde{N}V_{22} \end{bmatrix}, \quad \begin{bmatrix} U_{11}W & U_{12} \\ U_{21}W & U_{22} \end{bmatrix} = \begin{bmatrix} \tilde{W}V_{11} & \tilde{W}V_{12} \\ V_{21} & V_{22} \end{bmatrix}.$$

Comparing the entries of these matrices we find the relations $U_{12}N = V_{12}$ and $U_{12} = \tilde{W}V_{12}$, which lead to $U_{12} = \tilde{W}U_{12}N = \cdots = \tilde{W}^{\mu}U_{12}N^{\mu} = 0$. Analogously we derive $U_{21} = 0$. Then, the blocks $U_{11} = V_{11}$, $U_{22} = V_{22}$ must be nonsingular. It follows that

$$V_{11}W = \tilde{W}V_{11}, \quad V_{22}N = \tilde{N}V_{22}$$

holds true, that is, the matrices N and $\tilde{N}$ as well as W and $\tilde{W}$ are similar, and in particular, $\mu = \tilde{\mu}$ is valid. □

The real valued matrix N has the eigenvalue zero only, and can be transformed into its Jordan form by means of a real valued similarity transformation. Therefore, in Proposition 1.3, the transformation matrices L and K can be chosen such that N is in Jordan form.

Proposition 1.3 also holds true for complex valued matrices. This is a well-known result of Weierstraß and Kronecker, cf. [82]. The special pair given by (1.3) is said to be *Weierstraß–Kronecker form* of the original pair $\{E, F\}$.

Definition 1.4. The *Kronecker index* of a regular matrix pair $\{E, F\}$, $E, F \in L(\mathbb{R}^m)$, and the *Kronecker index of a regular DAE* (1.1) are defined to be the nilpotency order μ in the Weierstraß–Kronecker form (1.3). We write $\operatorname{ind}\{E, F\} = \mu$.

The Weierstraß–Kronecker form of a regular pair $\{E, F\}$ provides a broad insight into the structure of the associated DAE (1.1). Scaling of (1.1) by L and transforming $x = K \begin{bmatrix} y \\ z \end{bmatrix}$ leads to the equivalent decoupled system

$$y'(t) + Wy(t) = p(t), \quad t \in \mathcal{I}, \tag{1.4}$$
$$Nz'(t) + z(t) = r(t), \quad t \in \mathcal{I}, \tag{1.5}$$

with $Lq =: \begin{bmatrix} p \\ r \end{bmatrix}$. The first equation (1.4) represents a standard explicit ODE. The second one appears for $l > 0$, and it has the only solution

$$z(t) = \sum_{j=0}^{\mu-1} (-1)^j N^j r^{(j)}(t), \tag{1.6}$$

provided that r is smooth enough. The latter one becomes clear after recursive use of (1.5) since

$$z = r - Nz' = r - N(r - Nz')' = r - Nr' + N^2 z'' = r - Nr' + N^2(r - Nz')'' = \cdots$$

Expression (1.6) shows the dependence of the solution x on the derivatives of the source or perturbation term q. The higher the index μ, the more differentiations are involved. Only in the index-1 case do we have $N = 0$, hence $z(t) = r(t)$, and no derivatives are involved. Since numerical differentiations in these circumstances may cause considerable trouble, it is very important to know the index μ as well as details of the structure responsible for a higher index when modeling and simulating with DAEs in practice. The typical solution behavior of ill-posed problems can be observed in higher index DAEs: small perturbations of the right-hand side yield large changes in the solution. We demonstrate this by the next example.

Example 1.5 (Ill-posed behavior in case of a higher index DAE). The regular DAE

$$\underbrace{\begin{bmatrix} 1&0&0&0&0 \\ 0&0&1&0&0 \\ 0&0&0&1&0 \\ 0&0&0&0&1 \\ 0&0&0&0&0 \end{bmatrix}}_{E} x'(t) + \underbrace{\begin{bmatrix} -\alpha & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{F} x(t) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \gamma(t) \end{bmatrix},$$

completed by the initial condition $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix} x(0) = 0$, is uniquely solvable for each sufficiently smooth function γ. The identically zero solution corresponds to the vanishing input function $\gamma(t) = 0$. The solution corresponding to the small excitation $\gamma(t) = \varepsilon \frac{1}{n} \sin nt$, $n \in \mathbb{N}$, ε small, is

$$x_1(t) = \varepsilon \int_0^t n^2 e^{\alpha(t-s)} \cos ns \, \mathrm{d}s, \quad x_2(t) = \varepsilon n^2 \cos nt,$$
$$x_3(t) = -\varepsilon n \sin nt, \quad x_4(t) = -\varepsilon \cos nt, \quad x_5(t) = \varepsilon \frac{1}{n} \sin nt.$$

While the excitation tends to zero for $n \to \infty$, the first three solution components grow unboundedly. The solution value at $t = 0$,

$$x_1(0) = 0,\ x_2(0) = \varepsilon n^2,\ x_3(0) = 0,\ x_4(0) = -\varepsilon,\ x_5(0) = 0,$$

moves away from the origin with increasing n, and the origin is no longer a consistent value at $t = 0$ for the perturbed system, as it is the case for the unperturbed one. Figures 1.1 and 1.2 show γ and the response x_2 for $\varepsilon = 0.1$, $n = 1$ and $n = 100$. □

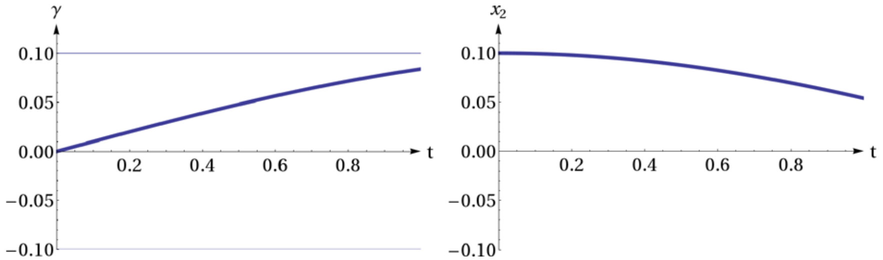


Fig. 1.1 γ and x_2 for $n = 1$

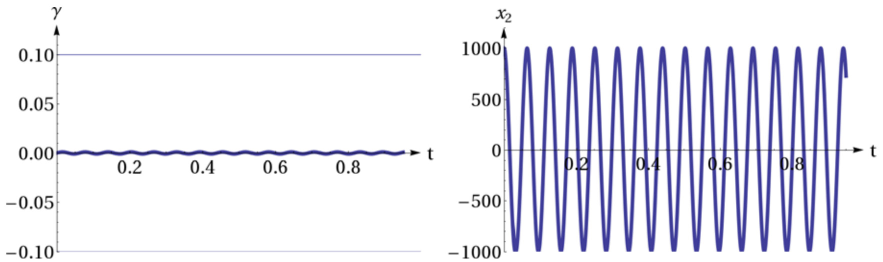


Fig. 1.2 γ and x_2 for $n = 100$

This last little constant coefficient example is relatively harmless. Time-dependent subspaces and nonlinear relations in more general DAEs may considerably amplify the bad behavior. For this reason one should be careful in view of numerical simulations. It may well happen that an integration code seemingly works, however it generates wrong results.

The general solution of a regular homogeneous DAE (1.2) is of the form

$$x(t) = K \begin{bmatrix} e^{-tW} \\ 0 \end{bmatrix} y_0,\ \ y_0 \in \mathbb{R}^{m-l}$$

which shows that the solution space has finite dimension $m - l$ and the solution depends smoothly on the initial value $y_0 \in \mathbb{R}^{m-l}$. Altogether, already for constant coefficient linear DAEs, the solutions feature an ambivalent behavior: they depend smoothly on certain initial values while they are ill-posed with respect to excitations.

The next theorem substantiates the above regularity notion.

Theorem 1.6. *The homogeneous DAE (1.2) has a finite-dimensional solution space if and only if the pair* $\{E,F\}$ *is regular.*

Proof. As we have seen before, if the pair $\{E,F\}$ is regular, then the solutions of (1.2) form an $(m-l)$-dimensional space. Conversely, let $\{E,F\}$ be a singular pair, i.e., $\det(\lambda E+F)\equiv 0$. For any set of $m+1$ different real values $\lambda_1,\dots,\lambda_{m+1}$ we find nontrivial vectors $\eta_1,\dots,\eta_{m+1}\in\mathbb{R}^m$ such that $(\lambda_i E+F)\eta_i=0$, $i=1,\dots,m+1$, and a nontrivial linear combination $\sum_{i=1}^{m+1}\alpha_i\eta_i=0$.

The function $x(t)=\sum_{i=1}^{m+1}\alpha_i e^{\lambda_i t}\eta_i$ does not vanish identically, and it satisfies the DAE (1.2) as well as the initial condition $x(0)=0$. For disjoint $(m+1)$-element sets $\{\eta_1,\dots,\eta_{m+1}\}$, one always has different solutions, and, consequently, the solution space of a homogeneous initial value problem (IVP) of (1.2) is not finite. □

Example 1.7 (Solutions of a nonregular DAE (cf. [97])). The pair $\{E,F\}$,

$$E=\begin{bmatrix}1&1&0&0\\0&0&0&1\\0&0&0&0\\0&0&1&0\end{bmatrix},\quad F=\begin{bmatrix}0&1&0&0\\0&0&0&0\\0&0&1&0\\0&0&0&0\end{bmatrix},\quad m=4,$$

is singular. In detail, the homogeneous DAE (1.2) reads

$$\begin{aligned}(x_1+x_2)' \quad + \quad x_2 &= 0,\\ x_4' \quad &= 0,\\ x_3 &= 0,\\ x_3' \quad &= 0.\end{aligned}$$

What does the solution space look like? Obviously, the component x_3 vanishes identically and x_4 is an arbitrary constant function. The remaining equation $(x_1+x_2)'+x_2=0$ is satisfied by any arbitrary continuous x_2, and the resulting expression for x_1 is

$$x_1(t)=c-x_2(t)-\int_0^t x_2(s)ds,$$

c being a further arbitrary constant. Clearly, this solution space does not have finite dimension, which confirms the assertion of Theorem 1.6. Indeed, the regularity assumption is violated since

$$p(\lambda)=\det(\lambda E+F)=\det\begin{bmatrix}\lambda&\lambda+1&0&0\\0&0&0&\lambda\\0&0&1&0\\0&0&\lambda&0\end{bmatrix}=0.$$

Notice that, in the case of nontrivial perturbations q, for the associated perturbed DAE (1.1) the consistency condition $q_3'=q_4$ must be valid for solvability. In practice, such unbalanced models should be avoided. However, in large dimensions m, this might not be a trivial task. □

We take a closer look at the subsystem (1.5) within the Weierstraß–Kronecker form, which is specified by the nilpotent matrix N. We may choose the transformation matrices L and K in such a way that N has Jordan form, say

$$N = \operatorname{diag}[J_1, \ldots, J_s], \tag{1.7}$$

with s nilpotent Jordan blocks

$$J_i = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix} \in L(\mathbb{R}^{k_i}),\ i = 1, \ldots, s,$$

where $k_1 + \cdots + k_s = l,\ \ \mu = \max\{k_i : i = 1, \ldots, s\}$. The Kronecker index μ equals the order of the maximal Jordan block in N.

The Jordan form (1.7) of N indicates the further decoupling of the subsystem (1.5) in accordance with the Jordan structure into s lower-dimensional equations

$$J_i \zeta_i'(t) + \zeta_i(t) = r_i(t),\ \ i = 1, \ldots, s.$$

We observe that $\zeta_{i,2}, \ldots, \zeta_{i,k_i}$ are components involved with derivatives whereas the derivative of the first component $\zeta_{i,1}$ is not involved. Notice that the value of $\zeta_{i,1}(t)$ depends on the $(k_i - 1)$-th derivative of $r_{i,k_i}(t)$ for all $i = 1, \ldots, s$ since

$$\zeta_{i,1}(t) = r_{i,1}(t) - \zeta_{i,2}'(t) = r_{i,1}(t) - r_{i,2}'(t) + \zeta_{i,3}'(t) = \cdots = \sum_{j=1}^{k_i} (-1)^{j-1} r_{i,j}^{(j-1)}(t).$$

1.2 Projector based decoupling of regular DAEs

1.2.1 Admissible matrix sequences and admissible projectors

Our aim is now a suitable rearrangement of terms within the equation

$$Ex'(t) + Fx(t) = q(t), \tag{1.8}$$

which allows for a similar insight into the structure of the DAE to that given by the Weierstraß–Kronecker form. However, we do not use transformations, but we work in terms of the original equation setting and apply a projector based decoupling concept. The construction is simple. We consider the DAE (1.8) with the coefficients $E, F \in L(\mathbb{R}^m)$.

Put $G_0 := E,\ B_0 := F,\ N_0 := \ker G_0$ and introduce $Q_0 \in L(\mathbb{R}^m)$ as a projector onto N_0. Let $P_0 := I - Q_0$ be the complementary one. Using the basic projector properties $Q_0^2 = Q_0$, $Q_0P_0 = P_0Q_0 = 0$, $P_0 + Q_0 = I$, $G_0Q_0 = 0$ and $G_0 = G_0P_0$ (see Appendix A), we rewrite the DAE (1.8) consecutively as

$$\begin{aligned}
& G_0x' + B_0x = q \\
\iff \quad & G_0P_0x' + B_0(Q_0 + P_0)x = q \\
\iff \quad & \underbrace{(G_0 + B_0Q_0)}_{=:G_1}(P_0x' + Q_0x) + \underbrace{B_0P_0}_{=:B_1}x = q \\
\iff \quad & G_1(P_0x' + Q_0x) + B_1x = q.
\end{aligned}$$

Next, let Q_1 be a projector onto $N_1 := \ker G_1$, and let $P_1 := I - Q_1$ the complementary one. We rearrange the last equation to

$$\begin{aligned}
& G_1P_1(P_0x' + Q_0x) + B_1(Q_1 + P_1)x = q \\
\iff \quad & \underbrace{(G_1 + B_1Q_1)}_{G_2}\big(P_1(P_0x' + Q_0x) + Q_1x\big) + \underbrace{B_1P_1}_{B_2}x = q
\end{aligned} \tag{1.9}$$

and so on. The goal is a matrix with maximal possible rank in front of the term containing the derivative x'.

We form, for $i \geq 0$,

$$G_{i+1} := G_i + B_iQ_i, \quad N_{i+1} := \ker G_{i+1}, \quad B_{i+1} := B_iP_i \tag{1.10}$$

and introduce $Q_{i+1} \in L(\mathbb{R}^m)$ as a projector onto N_{i+1} with $P_{i+1} := I - Q_{i+1}$. Denote $r_i := \operatorname{rank} G_i$ and introduce the product of projectors $\Pi_i := P_0 \cdots P_i$. These ranks and products of projectors will play a special role later on. From $B_{i+1} = B_iP_i = B_0\Pi_i$ we derive the inclusion $\ker \Pi_i \subseteq \ker B_{i+1}$ as an inherent property of our construction. Since $G_i = G_{i+1}P_i$, the further inclusions

$$\operatorname{im} G_0 \subseteq \operatorname{im} G_1 \subseteq \cdots \subseteq \operatorname{im} G_i \subseteq \operatorname{im} G_{i+1},$$

follow, and hence

$$r_0 \leq r_1 \leq \cdots \leq r_i \leq r_{i+1}.$$

An additional inherent property of the sequence (1.10) is given by

$$N_{i-1} \cap N_i \subseteq N_i \cap N_{i+1}, \quad i \geq 1. \tag{1.11}$$

Namely, if $G_{i-1}z = 0$ and $G_iz = 0$ are valid for a vector $z \in \mathbb{R}^m$, which corresponds to $P_{i-1}z = 0$ and $P_iz = 0$, i.e., $z = Q_iz$, then we can conclude that

$$G_{i+1}z = G_iz + B_iQ_iz = B_iz = B_{i-1}P_{i-1}z = 0.$$

From (1.11) we learn that a nontrivial intersection $N_{i_*-1} \cap N_{i_*}$ never allows an injective matrix G_i, $i > i_*$. As we will realize later (see Proposition 1.34), such a nontrivial intersection immediately indicates a singular matrix pencil $\lambda E + F$.

Again, we are aiming at a matrix G_κ the rank of which is as high as possible. However, how can we know whether the maximal rank has been reached? Appropriate criteria would be helpful. As we will see later on, for regular DAEs, the sequence terminates with a nonsingular matrix.

Example 1.8 (Sequence for a regular DAE). For the DAE

$$\begin{aligned} x_1' \quad + x_1 + x_2 + x_3 &= q_1, \\ x_3' \quad + x_2 \qquad &= q_2, \\ x_1 \qquad + x_3 &= q_3, \end{aligned}$$

the first matrices of our sequence are

$$G_0 = E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad B_0 = F = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

As a nullspace projector onto $\ker G_0$ we choose

$$Q_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and obtain } G_1 = G_0 + B_0 Q_0 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad B_1 = B_0 P_0 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Since G_1 is singular, we turn to the next level. We choose as a projector onto $\ker G_1$

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \text{ and arrive at } G_2 = G_1 + B_1 Q_1 = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 0 & 0 \end{bmatrix}.$$

The matrix G_2 is nonsingular, hence the maximal rank is reached and we stop constructing the sequence. Looking at the polynomial $p(\lambda) = \det(\lambda E + F) = 2\lambda$ we know this DAE to be regular. Later on we shall see that a nonsingular matrix G_2 is typical for regularity with Kronecker index 2. Observe further that the nullspaces N_0 and N_1 intersect trivially, and that the projector Q_1 is chosen such that $\Pi_0 Q_1 Q_0 = 0$ is valid, or equivalently, $N_0 \subseteq \ker \Pi_0 Q_1$. □

Example 1.9 (Sequence for a nonregular DAE). We consider the nonregular matrix pair from Example 1.7, that is

$$G_0 = E = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad B_0 = F = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Choosing

$$Q_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{yields} \quad G_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The matrix G_1 is singular. We turn to the next level. We pick

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{which implies} \quad G_2 = G_0.$$

We continue constructing

$$Q_{2j} = Q_0, \;\; G_{2j+1} = G_1, \;\; Q_{2j+1} = Q_1, \;\; G_{2j+2} = G_0, \;\; j \geq 1.$$

Here we have $r_i = 3$ for all $i \geq 0$. The maximal rank is already met by G_0, but there is no criterion which indicates this in time. Furthermore, $N_i \cap N_{i+1} = \{0\}$ holds true for all $i \geq 0$, such that there is no step indicating a singular pencil via property (1.11). Observe that the product $\Pi_0 Q_1 Q_0 = P_0 Q_1 Q_0$ does not vanish as it does in the previous example. □

The rather irritating experience with Example 1.9 leads us to the idea to refine the choice of the projectors by incorporating more information from the previous steps. So far, just the image spaces of the projectors Q_i are prescribed. We refine the construction by prescribing certain appropriate parts of their nullspaces, too. More precisely, we put parts of the previous nullspaces into the current one.

When constructing the sequence (1.10), we now proceed as follows. At any level we decompose

$$N_0 + \cdots + N_{i-1} = \widehat{N_i} \oplus X_i, \quad \widehat{N_i} := (N_0 + \cdots + N_{i-1}) \cap N_i, \tag{1.12}$$

where X_i is any complement to $\widehat{N_i}$ in $N_0 + \cdots + N_{i-1}$. We choose Q_i in such a way that the condition

$$X_i \subseteq \ker Q_i \tag{1.13}$$

is met. This is always possible since the subspaces $\widehat{N_i}$ and X_i intersect trivially (see Appendix, Lemma A.7). This restricts to some extent the choice of the projectors. However, a great variety of possible projectors is left. The choice (1.13) implies the projector products Π_i to be projectors again, cf. Proposition 1.13(2). Our structural analysis will significantly benefit from this property. We refer to Chapter 7 for a discussion of practical calculations.

If the intersection $\widehat{N_i} = (N_0 + \cdots + N_{i-1}) \cap N_i$ is trivial, then we have

$$X_i = N_0 + \cdots + N_{i-1} \subseteq \ker Q_i.$$

This is the case in Example 1.8 which shows a regular DAE.

Definition 1.10. For a given matrix pair $\{E,F\}, E,F \in L(\mathbb{R}^m)$, and an integer $\kappa \in \mathbb{N}$, we call the matrix sequence $G_0,\dots,G_\kappa$ an *admissible matrix sequence*, if it is built by the rule

Set $G_0 := E$, $B_0 := F$, $N_0 := \ker G_0$, and choose a projector $Q_0 \in L(\mathbb{R}^m)$ onto N_0.
For $i \geq 1$:

$$G_i := G_{i-1} + B_{i-1}Q_{i-1},$$
$$B_i := B_{i-1}P_{i-1}$$

$$N_i := \ker G_i, \quad \widehat{N}_i := (N_0 + \cdots + N_{i-1}) \cap N_i,$$

fix a complement X_i such that $N_0 + \cdots + N_{i-1} = \widehat{N}_i \oplus X_i$,
choose a projector Q_i such that $\operatorname{im} Q_i = N_i$ and $X_i \subseteq \ker Q_i$,
set $P_i := I - Q_i$, $\Pi_i := \Pi_{i-1}P_i$

The projectors $Q_0,\dots,Q_\kappa$ in an admissible matrix sequence are said to be *admissible*. The matrix sequence $G_0,\dots,G_\kappa$ is said to be *regular admissible*, if additionally,

$$\widehat{N}_i = \{0\}, \quad \forall\, i = 1,\dots,\kappa.$$

Then, also the projectors $Q_0,\dots,Q_\kappa$ are called *regular admissible*.

Admissible projectors are always cross-linked to the matrix function sequence. Changing a projector at a certain level the whole subsequent sequence changes accordingly. Later on we learn that nontrivial intersections $\widehat{N}_i$ indicate a singular matrix pencil.

The projectors in Example 1.8 are admissible but the projectors in Example 1.9 are not. We revisit Example 1.9 and provide admissible projectors.

Example 1.11 (Admissible projectors). Consider once again the singular pair from Examples 1.7 and 1.9. We start the sequence with the same matrices G_0, B_0, Q_0, G_1 as described in Example 1.9 but now we use an admissible projector Q_1. The nullspaces of G_0 and G_1 are given by

$$N_0 = \operatorname{span}\begin{bmatrix}1\\-1\\0\\0\end{bmatrix} \text{ and } N_1 = \operatorname{span}\begin{bmatrix}1\\0\\0\\0\end{bmatrix}.$$

This allows us to choose

$$Q_1 = \begin{bmatrix}1&1&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix},$$

which satisfies the condition $X_1 \subseteq \ker Q_1$, where $X_1 = N_0$ and $\widehat{N_1} = N_0 \cap N_1 = \{0\}$. It yields

$$G_2 = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Now we find $N_2 = \operatorname{span} \begin{bmatrix} -2 & 1 & 0 & 0 \end{bmatrix}^{\mathrm{T}}$ and with

$$N_0 + N_1 = N_0 \oplus N_1 = \operatorname{span}\left(\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right) = \operatorname{span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right),$$

we have $N_2 \subseteq N_0 + N_1$, $N_0 + N_1 + N_2 = N_0 + N_1$ as well as $\widehat{N_2} = (N_0 + N_1) \cap N_2 = N_2$. A possible complement X_2 to $\widehat{N_2}$ in $N_0 + N_1$ and an appropriate projector Q_2 are

$$X_2 = \operatorname{span} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 0 & -2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

This leads to $G_3 = G_2$, and the nontrivial intersection $N_2 \cap N_3$ indicates (cf. (1.11)) that also all further matrices G_i are singular. Proposition 1.34 below says that this indicates at the same time a singular matrix pencil. In the next steps, for $i \geq 3$, it follows that $N_i = N_2$ and $G_i = G_2$.

For admissible projectors Q_i, not only is their image $\operatorname{im} Q_i = N_i$ fixed, but also a part of $\ker Q_i$. However, there remains a great variety of possible projectors, since, except for the regular case, the subspaces X_i are not uniquely determined and further represent just a part of $\ker Q_i$. Of course, we could restrict the variety of projectors by prescribing special subspaces. For instance, we may exploit orthogonality as much as possible, which is favorable with respect to computational aspects.

Definition 1.12. The admissible projectors $Q_0, \ldots, Q_\kappa$ are called *widely orthogonal* if $Q_0 = Q_0^*$, and

$$X_i = \widehat{N_i}^{\perp} \cap (N_0 + \cdots + N_{i-1}), \tag{1.14}$$

as well as

$$\ker Q_i = [N_0 + \cdots + N_i]^{\perp} \oplus X_i, \quad i = 1, \ldots, \kappa, \tag{1.15}$$

hold true.

The widely orthogonal projectors are completely fixed and they have their advantages. However, in Subsection 2.2.3 we will see that it makes sense to work with sufficiently flexible admissible projectors.

The next assertions collect useful properties of admissible matrix sequences $G_0,\dots,G_\kappa$ and the associated admissible projectors $Q_0,\dots,Q_\kappa$ for a given pair $\{E,F\}$. In particular, the special role of the products $\Pi_i = P_0\cdots P_i$ is revealed. We emphasize this by using mainly the short notation Π_i.

Proposition 1.13. *Let $Q_0,\dots,Q_\kappa$ be admissible projectors for the pair $\{E,F\}$, $E,F\in L(\mathbb{R}^m)$. Then the following assertions hold true for $i=1,\dots,\kappa$:*

(1) $\ker\Pi_i = N_0+\cdots+N_i$.
(2) *The products $\Pi_i = P_0\cdots P_i$ and $\Pi_{i-1}Q_i = P_0\cdots P_{i-1}Q_i$, are again projectors.*
(3) $N_0+\cdots+N_{i-1}\subseteq\ker\Pi_{i-1}Q_i$.
(4) $B_i = B_i\Pi_{i-1}$.
(5) $\widehat{N}_i \subseteq N_i\cap\ker B_i = N_i\cap N_{i+1}\subseteq\widehat{N}_{i+1}$.
(6) *If $Q_0,\dots,Q_\kappa$ are widely orthogonal, then* $\operatorname{im}\Pi_i = [N_0+\cdots+N_i]^\perp$, $\Pi_i=\Pi_i^*$ *and* $\Pi_{i-1}Q_i = (\Pi_{i-1}Q_i)^*$.
(7) *If $Q_0,\dots,Q_\kappa$ are regular admissible, then* $\ker\Pi_{i-1}Q_i = \ker Q_i$ *and* $Q_iQ_j=0$ *for* $j=0,\dots,i-1$.

Proof. (1) ($\Rightarrow$) To show $\ker\Pi_i\subseteq N_0+\cdots+N_i$ for $i=1,\dots,\kappa$, we consider an element $z\in\ker\Pi_i$. Then,

$$0=\Pi_i z = P_0\cdots P_i z = \prod_{k=0}^{i}(I-Q_k)z.$$

Expanding the right-hand expression, we obtain

$$z=\sum_{k=0}^{i} Q_kH_kz \ \in N_0+\cdots+N_i$$

with suitable matrices H_k.
($\Leftarrow$) The other direction will be proven by induction. Starting the induction with $i=0$, we observe that $\ker\Pi_0=\ker P_0=N_0$. We suppose that $\ker\Pi_{i-1}=N_0+\cdots+N_{i-1}$ is valid. Because of

$$N_0+\cdots+N_i = X_i+\widehat{N}_i+N_i$$

each $z\in N_0+\cdots+N_i$ can be written as $z=x_i+\bar z_i+z_i$ with

$$x_i\in X_i\subseteq N_0+\cdots+N_{i-1}=\ker\Pi_{i-1},\quad \bar z_i\in\widehat{N}_i\subseteq N_i,\quad z_i\in N_i.$$

Since Q_i is admissible, we have $X_i\subseteq\ker Q_i$ and $N_i=\operatorname{im}Q_i$. Consequently,

$$\Pi_i z = \Pi_{i-1}(I-Q_i)z = \Pi_{i-1}(I-Q_i)x_i = \Pi_{i-1}x_i = 0$$

which implies $N_0+\cdots+N_i\subseteq\ker\Pi_i$ to be true.
(2) From (1) we know that $\operatorname{im}Q_j = N_j\subseteq\ker\Pi_i$ for $j\le i$. It follows that

$$\Pi_i P_j = \Pi_i(I - Q_j) = \Pi_i.$$

Consequently, $\Pi_i^2 = \Pi_i$ and $\Pi_i \Pi_{i-1} = \Pi_i$ imply

$$\begin{aligned}(\Pi_{i-1}Q_i)^2 = \Pi_{i-1}(I - P_i)\Pi_{i-1}Q_i &= \Pi_{i-1}Q_i - \Pi_i\Pi_{i-1}Q_i \\ &= \Pi_{i-1}Q_i - \Pi_i Q_i = \Pi_{i-1}Q_i.\end{aligned}$$

(3) For any $z \in N_0 + \cdots + N_{i-1}$, we know from (1) that $\Pi_{i-1}z = 0$ and $\Pi_i z = 0$. Thus

$$\Pi_{i-1}Q_i z = \Pi_{i-1}z - \Pi_i z = 0.$$

(4) By construction of B_i (see (1.10)), we find $B_i = B_0\Pi_{i-1}$. Using (2), we get that

$$B_i = B_0\Pi_{i-1} = B_0\Pi_{i-1}\Pi_{i-1} = B_i\Pi_{i-1}.$$

(5) First, we show that $\widehat{N_i} \subseteq N_i \cap \ker B_i$. For $z \in \widehat{N_i} = (N_0 + \cdots + N_{i-1}) \cap N_i$ we find $\Pi_{i-1}z = 0$ from (1) and, hence, $B_i z = B_0\Pi_{i-1}z = 0$ using (4). Next,

$$N_i \cap \ker B_i = N_i \cap N_{i+1}$$

since $G_{i+1}z = (G_i + B_iQ_i)z = B_i z$ for any $z \in N_i = \operatorname{im} Q_i = \ker G_i$. Finally,

$$\widehat{N}_{i+1} = (N_0 + \cdots + N_i) \cap N_{i+1} \text{ implies immediately that } N_i \cap N_{i+1} \subseteq \widehat{N}_{i+1}.$$

(6) We use induction to show that $\operatorname{im}\Pi_i = [N_0 + \cdots + N_i]^\perp$. Starting with $i = 0$, we know that $\operatorname{im}\Pi_0 = N_0^\perp$ since $Q_0 = Q_0^*$.
Since $X_i \subseteq N_0 + \cdots + N_{i-1}$ (see (1.14)) we derive from (1) that $\Pi_{i-1}X_i = 0$. Regarding (1.15), we find

$$\operatorname{im}\Pi_i = \Pi_{i-1}\operatorname{im}P_i = \Pi_{i-1}([N_0 + \cdots + N_i]^\perp + X_i) = \Pi_{i-1}([N_0 + \cdots + N_i]^\perp).$$

Using $[N_0 + \cdots + N_i]^\perp \subseteq [N_0 + \cdots + N_{i-1}]^\perp = \operatorname{im}\Pi_{i-1}$ we conclude

$$\operatorname{im}\Pi_i = \Pi_{i-1}([N_0 + \cdots + N_i])^\perp = [N_0 + \cdots + N_i]^\perp.$$

In consequence, Π_i is the orthoprojector onto $[N_0 + \cdots + N_i]^\perp$ along $N_0 + \cdots + N_i$, i.e., $\Pi_i = \Pi_i^*$. It follows that

$$\Pi_{i-1}Q_i = \Pi_{i-1} - \Pi_i = \Pi_{i-1}^* - \Pi_i^* = (\Pi_{i-1} - \Pi_i)^* = (\Pi_{i-1}Q_i)^*.$$

(7) Let $\widehat{N_i} = 0$ be valid. Then, $X_i = N_0 + \cdots + N_{i-1} = N_0 \oplus \cdots \oplus N_{i-1}$ and, therefore,

$$\ker\Pi_{i-1} \overset{(1)}{=} N_0 \oplus \cdots \oplus N_{i-1} = X_i \subseteq \ker Q_i.$$

This implies $Q_iQ_j = 0$ for $j = 0, \ldots, i-1$. Furthermore, for any $z \in \ker\Pi_{i-1}Q_i$, we have $Q_i z \in \ker\Pi_{i-1} \subseteq \ker Q_i$, which means that $z \in \ker Q_i$.

□

Remark 1.14. If the projectors $Q_0,\dots,Q_\kappa$ are regular admissible, and the $\Pi_0,\dots,\Pi_\kappa$ are symmetric, then $Q_0,\dots,Q_\kappa$ are widely orthogonal. This is a consequence of the properties

$$\operatorname{im}\Pi_i = (\ker\Pi_i)^\perp = (N_0\oplus\cdots\oplus N_i)^\perp, \quad \ker Q_i = \operatorname{im}\Pi_i\oplus X_i \quad \text{for } i=1,\dots,\kappa.$$

In more general cases, if there are nontrivial intersections $\widehat{N_i}$, widely orthogonal projectors are given, if the Π_i are symmetric and, additionally, the conditions $Q_i\Pi_i = 0$, $P_i(I-\Pi_{i-1}) = (P_i(I-\Pi_{i-1}))^*$ are valid (cf. Chapter 7).

Now we are in a position to provide a result which plays a central role in the projector approach of regular DAEs.

Theorem 1.15. *If, for the matrix pair $\{E,F\}$, $E,F\in L(\mathbb{R}^m)$, an admissible matrix sequence $(G_i)_{i\geq 0}$ contains an integer μ such that G_μ is nonsingular, then the representations*

$$G_\mu^{-1}E = \Pi_{\mu-1} + (I-\Pi_{\mu-1})G_\mu^{-1}E(I-\Pi_{\mu-1}) \tag{1.16}$$

$$G_\mu^{-1}F = Q_0+\cdots+Q_{\mu-1} + (I-\Pi_{\mu-1})G_\mu^{-1}F\Pi_{\mu-1} + \Pi_{\mu-1}G_\mu^{-1}F\Pi_{\mu-1} \tag{1.17}$$

are valid and $\{E,F\}$ is a regular pair.

Proof. Let G_μ be nonsingular. Owing to Proposition 1.13 we express

$$\begin{aligned} F(I-\Pi_{\mu-1}) &= F(Q_0+\Pi_0Q_1+\cdots+\Pi_{\mu-2}Q_{\mu-1}) \\ &= B_0Q_0+B_1Q_1+\cdots+B_{\mu-1}Q_{\mu-1} \\ &= G_\mu Q_0+G_\mu Q_1+\cdots+G_\mu Q_{\mu-1} \\ &= G_\mu(Q_0+Q_1+\cdots+Q_{\mu-1}), \end{aligned}$$

therefore

$$\Pi_{\mu-1}G_\mu^{-1}F(I-\Pi_{\mu-1}) = 0. \tag{1.18}$$

Additionally, we have $G_\mu = E+F(I-\Pi_{\mu-1})$, thus $I = G_\mu^{-1}E + G_\mu^{-1}F(I-\Pi_{\mu-1})$ and $\Pi_{\mu-1} = \Pi_{\mu-1}G_\mu^{-1}E = G_\mu^{-1}E\Pi_{\mu-1}$. From these properties it follows that

$$\Pi_{\mu-1}G_\mu^{-1}E(I-\Pi_{\mu-1}) = 0, \tag{1.19}$$

which proves the expressions (1.16), (1.17).

Denote the finite set consisting of all eigenvalues of the matrix $-\Pi_{\mu-1}G_\mu^{-1}F$ by Λ. We show the matrix $\lambda E+F$ to be nonsingular for each arbitrary λ not belonging to Λ, which proves the matrix pencil to be regular. The equation $(\lambda E+F)z=0$ is equivalent to

$$\begin{aligned} &\lambda G_\mu^{-1}Ez + G_\mu^{-1}Fz = 0 \quad \Longleftrightarrow \\ &\lambda G_\mu^{-1}E\Pi_{\mu-1}z + \lambda G_\mu^{-1}E(I-\Pi_{\mu-1})z + G_\mu^{-1}F\Pi_{\mu-1}z + G_\mu^{-1}F(I-\Pi_{\mu-1})z = 0 \end{aligned} \tag{1.20}$$

Multiplying (1.20) by $\Pi_{\mu-1}$ and regarding (1.18)–(1.19), yields

$$\lambda \Pi_{\mu-1} z + \Pi_{\mu-1} G_\mu^{-1} F \Pi_{\mu-1} z = (\lambda I + \Pi_{\mu-1} G_\mu^{-1} F) \Pi_{\mu-1} z = 0,$$

which implies $\Pi_{\mu-1} z = 0$ for $\lambda \notin \Lambda$. Using $\Pi_{\mu-1} z = 0$, equation (1.20) multiplied by $I - \Pi_{\mu-1}$ reduces to

$$\lambda (I - \Pi_{\mu-1}) G_\mu^{-1} E (I - \Pi_{\mu-1}) z + (I - \Pi_{\mu-1}) G_\mu^{-1} F (I - \Pi_{\mu-1}) z = 0.$$

Replacing $G_\mu^{-1} E = I - G_\mu^{-1} F (I - \Pi_{\mu-1})$ we find

$$\lambda (I - \Pi_{\mu-1}) z + (1 - \lambda)(I - \Pi_{\mu-1}) G_\mu^{-1} F (I - \Pi_{\mu-1})(I - \Pi_{\mu-1}) z = 0.$$

If $\lambda = 1$ then this immediately implies $z = 0$. If $\lambda \neq 1$ it holds that

$$\Big(\frac{\lambda}{1-\lambda} I + \underbrace{(I - \Pi_{\mu-1}) G_\mu^{-1} F (I - \Pi_{\mu-1})}_{Q_0 + \cdots + Q_{\mu-1}} \Big) \underbrace{(I - \Pi_{\mu-1}) z}_{z} = 0.$$

Multiplication by $Q_{\mu-1}$ gives $Q_{\mu-1} z = 0$. Then multiplication by $Q_{\mu-2}$ yields $Q_{\mu-2} z = 0$, and so on. Finally we obtain $Q_0 z = 0$ and hence $z = (I - \Pi_{\mu-1}) z = Q_0 z + \cdots + \Pi_{\mu-2} Q_{\mu-1} z = 0$. □

Once more we emphasize that the matrix sequence depends on the choice of the admissible projectors. However, the properties that are important later on are independent of the choice of the projectors, as the following theorem shows.

Theorem 1.16. *For any pair $\{E, F\}$, $E, F \in L(\mathbb{R}^m)$, the subspaces $N_0 + \cdots + N_i$, $\widehat{N_i}$ and $\operatorname{im} G_i$ are independent of the special choice of the involved admissible projectors.*

Proof. All claimed properties are direct and obvious consequences of Lemma 1.18 below. □

Theorem 1.16 justifies the next definition.

Definition 1.17. For each arbitrary matrix pair $\{E, F\}$, $E, F \in L(\mathbb{R}^m)$, the integers $r_i := \operatorname{rank} G_i$, $i \geq 0$, $u_i := \dim \widehat{N_i}$ $i \geq 1$, which arise from an admissible matrix sequence $(G_i)_{i \geq 0}$, are called *structural characteristic values*.

Lemma 1.18. *Let $Q_0, \ldots, Q_\kappa$ and $\bar{Q}_0, \ldots, \bar{Q}_\kappa$ be any two admissible projector sequences for the pair $\{E, F\}$, $E, F \in L(\mathbb{R}^m)$, and N_j, $\bar{N}_j$, G_j, $\bar{G}_j$, etc. the corresponding subspaces and matrices. Then it holds that:*

(1) $\bar{N}_0 + \cdots + \bar{N}_j = N_0 + \cdots + N_j$, *for* $j = 0, \ldots, \kappa$.

(2) $\bar{G}_j = G_j Z_j$, $\bar{B}_j = B_j + G_j \sum_{l=0}^{j-1} Q_l \mathfrak{A}_{jl}$, *for* $j = 0, \ldots, \kappa$,
with nonsingular matrices $Z_0, \ldots, Z_{\kappa+1}$ given by $Z_0 := I$, $Z_{j+1} := Y_{j+1} Z_j$,

$$Y_1 := I + Q_0(\bar{Q}_0 - Q_0) = I + Q_0\bar{Q}_0 P_0,$$
$$Y_{j+1} := I + Q_j(\bar{\Pi}_{j-1}\bar{Q}_j - \Pi_{j-1}Q_j) + \sum_{l=0}^{j-1} Q_l \mathfrak{A}_{jl}\bar{Q}_j,$$

where $\mathfrak{A}_{jl} = \bar{\Pi}_{j-1}$ *for* $l = 0, \ldots, j-1$.

(3) $\bar{G}_{\kappa+1} = G_{\kappa+1}Z_{\kappa+1}$ *and* $\bar{N}_0 + \cdots + \bar{N}_{\kappa+1} = N_0 + \cdots + N_{\kappa+1}$.

(4) $(\bar{N}_0 + \cdots + \bar{N}_{j-1}) \cap \bar{N}_j = (N_0 + \cdots + N_{j-1}) \cap N_j$ *for* $j = 1, \ldots, \kappa+1$.

Remark 1.19. The introduction of $\mathfrak{A}_{il}$ seems to be unnecessary at this point. We use these extra terms to emphasize the great analogy to the case of DAEs with time-dependent coefficients (see Lemma 2.12). The only difference between both cases is given in the much more elaborate representation of $\mathfrak{A}_{il}$ for time-dependent coefficients.

Proof. We prove (1) and (2) together by induction. For $i = 0$ we have

$$\bar{G}_0 = E = G_0, \quad \bar{B}_0 = F = B_0, \quad \bar{N}_0 = \ker \bar{G}_0 = \ker G_0 = N_0, \quad Z_0 = I.$$

To apply induction we suppose the relations

$$\bar{N}_0 + \cdots + \bar{N}_j = N_0 + \cdots + N_j, \tag{1.21}$$

$$\bar{G}_j = G_j Z_j, \;\; \bar{B}_j = B_j + G_j \sum_{l=0}^{j-1} Q_l \mathfrak{A}_{jl} \tag{1.22}$$

to be valid for $j \le i$ with nonsingular Z_j as described above, and

$$Z_j^{-1} = I + \sum_{l=0}^{j-1} Q_l \mathfrak{C}_{jl}$$

with certain $\mathfrak{C}_{jl}$. Comparing $\bar{G}_{i+1}$ and G_{i+1} we write

$$\bar{G}_{i+1} = \bar{G}_i + \bar{B}_i\bar{Q}_i = G_i Z_i + \bar{B}_i\bar{Q}_i Z_i + \bar{B}_i\bar{Q}_i(I - Z_i) \tag{1.23}$$

and consider the last term in more detail. We have, due to the form of Y_l, induction assumption (1.21) and $\operatorname{im}(Y_j - I) \subseteq N_0 + \cdots + N_{j-1} = \ker \Pi_{j-1}$ given for all $j \ge 0$ (see Proposition 1.13) that

$$N_0 + \cdots + N_{j-1} \subseteq \ker \Pi_{j-1}Q_j, \quad \bar{N}_0 + \cdots + \bar{N}_{j-1} \subseteq \ker \bar{\Pi}_{j-1}\bar{Q}_j, \quad j \le i, \tag{1.24}$$

and therefore,

$$Y_{j+1} - I = (Y_{j+1} - I)\Pi_{j-1}, \quad j = 1, \ldots, i. \tag{1.25}$$

This implies

$$\operatorname{im}(Y_j - I) \subseteq \ker(Y_{j+1} - I), \quad j = 1, \ldots, i. \tag{1.26}$$

Concerning $Z_j = Y_j Z_{j-1}$ and using (1.26), a simple induction proof shows

$$Z_j - I = \sum_{l=1}^{j} (Y_l - I), \quad j = 1, \ldots, i,$$

to be satisfied. Consequently,

$$\operatorname{im}(I - Z_i) \subseteq N_0 + \cdots + N_{i-1} = \bar{N}_0 + \cdots + \bar{N}_{i-1} \subseteq \ker \bar{Q}_i.$$

Using (1.23), we get

$$\bar{G}_{i+1} = G_i Z_i + \bar{B}_i \bar{Q}_i Z_i,$$

which leads to

$$\bar{G}_{i+1} Z_i^{-1} = G_i + \bar{B}_i \bar{Q}_i = G_i + B_i Q_i + (\bar{B}_i \bar{Q}_i - B_i Q_i).$$

We apply the induction assumption (1.22) to find

$$\bar{G}_{i+1} Z_i^{-1} = G_{i+1} + B_i(\bar{Q}_i - Q_i) + G_i \sum_{l=0}^{i-1} Q_l \mathfrak{A}_{il} \bar{Q}_i.$$

Induction assumption (1.21) and Proposition 1.13 imply $\ker \bar{\Pi}_{i-1} = \ker \Pi_{i-1}$ and hence

$$B_i = B_0 \Pi_{i-1} = B_0 \Pi_{i-1} \bar{\Pi}_{i-1} = B_i \bar{\Pi}_{i-1}.$$

Finally,

$$\begin{aligned} \bar{G}_{i+1} Z_i^{-1} &= G_{i+1} + B_i(\bar{\Pi}_{i-1} \bar{Q}_i - \Pi_{i-1} Q_i) + G_{i+1} \sum_{l=0}^{i-1} Q_l \mathfrak{A}_{il} \bar{Q}_i \\ &= G_{i+1} + B_i Q_i (\bar{\Pi}_{i-1} \bar{Q}_i - \Pi_{i-1} Q_i) + G_{i+1} \sum_{l=0}^{i-1} Q_l \mathfrak{A}_{il} \bar{Q}_i \; = \; G_{i+1} Y_{i+1}, \end{aligned}$$

which means that

$$\bar{G}_{i+1} = G_{i+1} Y_{i+1} Z_i = G_{i+1} Z_{i+1}. \tag{1.27}$$

Next, we will show Z_{i+1} to be nonsingular. Owing to the induction assumption, we know that Z_i is nonsingular. Considering the definition of Z_{i+1} we have to show Y_{i+1} to be nonsingular. Firstly,

$$\Pi_i Y_{i+1} = \Pi_i \tag{1.28}$$

since $\operatorname{im} Q_j \subseteq \ker \Pi_i$ for $j \leq i$. This follows immediately from the definition of Y_{i+1} and Proposition 1.13 (1). Using the induction assumption (1.21), Proposition 1.13 and Lemma A.3, we find

$$\Pi_j \bar{\Pi}_j = \Pi_j, \quad \bar{\Pi}_j \Pi_j = \bar{\Pi}_j \quad \text{and} \quad \Pi_j \Pi_j = \Pi_j \qquad \text{for } j = 0, \ldots, i.$$

This implies that

$$\Pi_{i-1}(Y_{i+1} - I) = \Pi_{i-1}(Y_{i+1} - I)\Pi_i \tag{1.29}$$

because

$$\begin{aligned}
\Pi_{i-1}(Y_{i+1}-I) &\overset{\text{Prop. 1.13(1)}}{=} \Pi_{i-1}Q_i(\bar{\Pi}_{i-1}\bar{Q}_i-\Pi_{i-1}Q_i) \\
&= (\Pi_{i-1}-\Pi_i)(\bar{\Pi}_{i-1}\bar{Q}_i-\Pi_{i-1}Q_i) \\
&= \Pi_{i-1}(\bar{Q}_i-Q_i) \quad = \Pi_{i-1}(P_i-\bar{P}_i) \\
&= \Pi_i-\Pi_{i-1}\bar{\Pi}_{i-1}\bar{P}_i = \Pi_i-\Pi_{i-1}\bar{\Pi}_i \\
&= \Pi_i-\Pi_{i-1}\bar{\Pi}_i\Pi_i \quad = (I-\Pi_{i-1}\bar{\Pi}_i)\Pi_i.
\end{aligned}$$

Equations (1.28) and (1.29) imply

$$\Pi_{i-1}(Y_{i+1}-I)=\Pi_{i-1}(Y_{i+1}-I)\Pi_i=\Pi_{i-1}(Y_{i+1}-I)\Pi_iY_{i+1}$$

and, consequently,

$$\begin{aligned}
I = Y_{i+1}-(Y_{i+1}-I) &\overset{(1.25)}{=} Y_{i+1}-(Y_{i+1}-I)\Pi_{i-1} \\
&= Y_{i+1}-(Y_{i+1}-I)\Pi_{i-1}\{(I-\Pi_{i-1})Y_{i+1}+\Pi_{i-1}\} \\
&= Y_{i+1}-(Y_{i+1}-I)\Pi_{i-1}\{Y_{i+1}-\Pi_{i-1}(Y_{i+1}-I)\} \\
&= Y_{i+1}-(Y_{i+1}-I)\Pi_{i-1}\{Y_{i+1}-\Pi_{i-1}(Y_{i+1}-I)\Pi_iY_{i+1}\} \\
&= (I-(Y_{i+1}-I)\{I-\Pi_{i-1}(Y_{i+1}-I)\Pi_i\})Y_{i+1}.
\end{aligned}$$

This means that Y_{i+1} is nonsingular and

$$Y_{i+1}^{-1}=I-(Y_{i+1}-I)\{I-\Pi_{i-1}(Y_{i+1}-I)\Pi_i\}.$$

Then also $Z_{i+1}=Y_{i+1}Z_i$ is nonsingular, and

$$Z_{i+1}^{-1}=Z_i^{-1}Y_{i+1}^{-1}=(I+\sum_{l=0}^{i-1}Q_l\mathfrak{C}_{il})Y_{i+1}^{-1}=I+\sum_{l=0}^{i}Q_l\mathfrak{C}_{i+1\,l}$$

with certain coefficients $\mathfrak{C}_{i+1\,l}$. From (1.27) we conclude $\bar{N}_{i+1}=Z_{i+1}^{-1}N_{i+1}$, and, due to the special form of Z_{i+1}^{-1},

$$\bar{N}_{i+1}\subseteq N_0+\cdots+N_{i+1},\ \bar{N}_0+\cdots+\bar{N}_{i+1}\subseteq N_0+\cdots+N_{i+1}.$$

Owing to the property $\operatorname{im}(Z_{i+1}-I)\subseteq N_0+\cdots+N_i=\bar{N}_0+\cdots+\bar{N}_i$, it holds that

$$N_{i+1}=Z_{i+1}\bar{N}_{i+1}=(I+(Z_{i+1}-I))\bar{N}_{i+1}\subseteq\bar{N}_0+\cdots+\bar{N}_{i+1}.$$

Thus, $N_0+\cdots+N_{i+1}\subseteq\bar{N}_0+\cdots+\bar{N}_{i+1}$ is valid. For symmetry reasons we have

$$N_0+\cdots+N_{i+1}=\bar{N}_0+\cdots+\bar{N}_{i+1}.$$

Finally, we derive from the induction assumption that

$$\begin{aligned}
\bar{B}_{i+1} &= \bar{B}_i\bar{P}_i = \big(B_i + G_i\sum_{l=0}^{i-1} Q_l\mathfrak{A}_{il}\big)\bar{P}_i \\
&= B_iP_i\bar{P}_i + B_iQ_i\bar{P}_i + G_{i+1}\sum_{l=0}^{i-1} Q_l\mathfrak{A}_{il}\bar{P}_i \\
&= B_iP_i + B_iQ_i\bar{\Pi}_i + G_{i+1}\sum_{l=0}^{i-1} Q_l\mathfrak{A}_{il}\bar{P}_i \;=\; B_{i+1} + G_{i+1}\sum_{l=0}^{i} Q_l\mathfrak{A}_{i+1,l}
\end{aligned}$$

with $\mathfrak{A}_{i+1,l} = \mathfrak{A}_{il}\bar{P}_i$, $l = 0,\dots,i-1$, $\mathfrak{A}_{i+1,l} = \bar{\Pi}_i$, and therefore, for $l \le i-1$,

$$\mathfrak{A}_{i+1,l} = \mathfrak{A}_{il}\bar{P}_i = \mathfrak{A}_{i-1,l}\bar{P}_{i-1}\bar{P}_i = \mathfrak{A}_{l+1,l}\bar{P}_{l+1}\cdots\bar{P}_i = \bar{\Pi}_l\bar{P}_{l+1}\cdots\bar{P}_i = \bar{\Pi}_i.$$

We have proved assertions (1) and (2), and (3) is a simple consequence. Next we prove assertion (4). By assertion (1) from Lemma 1.13, we have $N_0 + \cdots + N_i = \ker\Pi_i$ and

$$\begin{aligned}
G_{i+1} &= G_0 + B_0Q_0 + \cdots + B_iQ_i = G_0 + B_0Q_0 + B_1P_0Q_1 + \cdots + B_iP_0\cdots P_{i-1}Q_i \\
&= G_0 + B_0(Q_0 + P_0Q_1 + \cdots + P_0\cdots P_{i-1}Q_i) \\
&= G_0 + B_0(I - P_0\cdots P_i) = G_0 + B_0(I - \Pi_i).
\end{aligned}$$

This leads to the description

$$\begin{aligned}
\widehat{N}_{i+1} = (N_0 + \cdots + N_i)\cap N_{i+1} &= \{z\in\mathbb{R}^m : \Pi_i z = 0,\; G_0z + B_0(I-\Pi_i)z = 0\} \\
&= \{z\in\mathbb{R}^m : z\in N_0 + \cdots + N_i,\; G_0z + B_0z = 0\} \\
&= \{z\in\mathbb{R}^m : z\in \bar{N}_0 + \cdots + \bar{N}_i,\; \bar{G}_0z + \bar{B}_0z = 0\} \\
&= (\bar{N}_0 + \cdots + \bar{N}_i)\cap\bar{N}_{i+1}.
\end{aligned}$$

□

1.2.2 Decoupling by admissible projectors

In this subsection we deal with matrix pairs $\{E,F\}$, $E,F \in L(\mathbb{R}^m)$, the admissible matrix sequence $(G_i)_{i\ge 0}$ of which reaches a nonsingular matrix G_μ. Those matrix pairs as well as the associated DAEs

$$Ex'(t) + Fx(t) = q(t) \tag{1.30}$$

are regular by Theorem 1.15. They have the structural characteristic values

$$r_0 \le \cdots \le r_{\mu-1} < r_\mu = m.$$

The nonsingular matrix G_μ allows for a projector based decoupling such that the decoupled version of the given DAE looks quite similar to the Weierstraß–Kronecker form.

We stress that, at the same time, our discussion should serve as a model for a corresponding decoupling of time-dependent linear DAEs for which we do not have a Weierstraß–Kronecker form.

When constructing an admissible matrix function sequence $(G_i)_{i\geq 0}$ we have in mind a *rearrangement of terms within the original DAE* (1.30) such that the solution components $\Pi_{\mu-1}x(t)$ and $(I-\Pi_{\mu-1})x(t)$ are separated as far as possible and the nonsingular matrix G_μ occurs in front of the derivative $(\Pi_{\mu-1}x(t))'$. Let the admissible matrix sequence (Definition 1.10) starting from $G_0 = E, B_0 = F$ be realized up to G_μ which is nonsingular. Let $\mu \in \mathbb{N}$ be the smallest such index.

Consider the involved admissible projectors $Q_0,\dots,Q_\mu$. We have $Q_\mu = 0, P_\mu = I$, $\Pi_\mu = \Pi_{\mu-1}$ for trivial reasons. Due to Proposition 1.13, the intersections $\widehat{N}_i$ are trivial,

$$\widehat{N}_i = N_i \cap (N_0 + \cdots + N_{i-1}) = \{0\}, \quad i = 1,\dots,\mu-1,$$

and therefore

$$N_0 + \cdots + N_{i-1} = N_0 \oplus \cdots \oplus N_{i-1}, \quad X_i = N_0 \oplus \cdots \oplus N_{i-1}, \quad i = 1,\dots,\mu-1. \quad (1.31)$$

From (1.31) we derive the relations

$$Q_i Q_j = 0, \quad j = 0,\dots,i-1, \quad i = 1,\dots,\mu-1, \quad (1.32)$$

which are very helpful in computations. Recall the properties

$$\begin{aligned} G_i P_{i-1} &= G_{i-1}, \quad B_i = B_i \Pi_{i-1}, \; i = 1,\dots,\mu, \\ G_i Q_j &= B_j Q_j, \; j = 0,\dots,i-1, \quad i = 0,\dots,\mu-1, \end{aligned}$$

from Section 1.2 which will be used frequently.

Applying $G_0 = G_0 P_0 = G_0 \Pi_0$ we rewrite the DAE (1.30) as

$$G_0(\Pi_0 x(t))' + B_0 x(t) = q(t), \quad (1.33)$$

and then, with $B_0 = B_0 P_0 + B_0 Q_0 = B_0 \Pi_0 + G_1 Q_0$, as

$$G_1 P_1 P_0 (\Pi_0 x(t))' + B_0 \Pi_0 x(t) + G_1 Q_0 x(t) = q(t).$$

Now we use the relation

$$\begin{aligned} G_1 P_1 P_0 &= G_1 \Pi_0 P_1 P_0 + G_1 (I - \Pi_0) P_1 P_0 \\ &= G_1 \Pi_1 - G_1 (I - \Pi_0) Q_1 \\ &= G_1 \Pi_1 - G_1 (I - \Pi_0) Q_1 \Pi_0 Q_1 \end{aligned}$$

to replace the first term. This yields

$$G_1(\Pi_1 x(t))' + B_1 x(t) + G_1\{Q_0 x(t) - (I - \Pi_0) Q_1 (\Pi_0 Q_1 x(t))'\} = q(t).$$

Proceeding further by induction we suppose

$$G_i(\Pi_i x(t))' + B_i x(t) + G_i \sum_{l=0}^{i-1} \{Q_l x(t) - (I-\Pi_l)Q_{l+1}(\Pi_l Q_{l+1}x(t))'\} = q(t) \quad (1.34)$$

and, in the next step, using the properties $G_{i+1}P_{i+1}P_i = G_i$, $B_i Q_i = G_{i+1}Q_i$, $G_i Q_l = G_{i+1}Q_l$, $l = 0,\dots,i-1$, and

$$\begin{aligned} P_{i+1}P_i\Pi_i &= \Pi_i P_{i+1}P_i\Pi_i + (I-\Pi_i)P_{i+1}P_i\Pi_i \\ &= \Pi_{i+1} - (I-\Pi_i)Q_{i+1} \\ &= \Pi_{i+1} - (I-\Pi_i)Q_{i+1}\Pi_i Q_{i+1}, \end{aligned}$$

we reach

$$G_{i+1}(\Pi_{i+1}x(t))' + B_{i+1}x(t) + G_{i+1}\sum_{l=0}^{i} \{Q_l x(t) - (I-\Pi_l)Q_{l+1}(\Pi_l Q_{l+1}x(t))'\} = q(t),$$

so that expression (1.34) can be used for all $i = 1,\dots,\mu$. In particular, we obtain

$$G_\mu(\Pi_\mu x(t))' + B_\mu x(t) + G_\mu \sum_{l=0}^{\mu-1} \{Q_l x(t) - (I-\Pi_l)Q_{l+1}(\Pi_l Q_{l+1}x(t))'\} = q(t). \quad (1.35)$$

Taking into account that $Q_\mu = 0$, $P_\mu = I$, $\Pi_\mu = \Pi_{\mu-1}$, and scaling with G_μ^{-1} we derive the equation

$$(\Pi_{\mu-1}x(t))' + G_\mu^{-1}B_\mu x(t) + \sum_{l=0}^{\mu-1} Q_l x(t) - \sum_{l=0}^{\mu-2} (I-\Pi_l)Q_{l+1}(\Pi_l Q_{l+1}x(t))' = G_\mu^{-1}q(t). \quad (1.36)$$

In turn, equation (1.36) can be decoupled into two parts, the explicit ODE with respect to $\Pi_{\mu-1}x(t)$,

$$(\Pi_{\mu-1}x(t))' + \Pi_{\mu-1}G_\mu^{-1}B_\mu x(t) = \Pi_{\mu-1}G_\mu^{-1}q(t), \quad (1.37)$$

and the remaining equation

$$(I-\Pi_{\mu-1})G_\mu^{-1}B_\mu x(t) + \sum_{l=0}^{\mu-1} Q_l x(t) - \sum_{l=0}^{\mu-2} (I-\Pi_l)Q_{l+1}(\Pi_l Q_{l+1}x(t))' = (I-\Pi_{\mu-1})G_\mu^{-1}q(t). \quad (1.38)$$

Next, we show that equation (1.38) uniquely defines the component $(I-\Pi_{\mu-1})x(t)$ in terms of $\Pi_{\mu-1}x(t)$. We decouple equation (1.38) once again into μ further parts according to the decomposition

$$I-\Pi_{\mu-1} = Q_0P_1\cdots P_{\mu-1} + Q_1P_2\cdots P_{\mu-1} + \cdots + Q_{\mu-2}P_{\mu-1} + Q_{\mu-1}. \tag{1.39}$$

Notice that $Q_iP_{i+1}\cdots P_{\mu-1}$, $i=0,\ldots,\mu-2$ are projectors, too, and

$$\begin{aligned} Q_iP_{i+1}\cdots P_{\mu-1}Q_i &= Q_i, \\ Q_iP_{i+1}\cdots P_{\mu-1}Q_j &= 0, \quad \text{if } i\neq j, \\ Q_iP_{i+1}\cdots P_{\mu-1}(I-\Pi_l)Q_{l+1} &= Q_i(I-\Pi_l)Q_{l+1} = 0, \quad \text{for } l=0,\ldots,i-1, \\ Q_iP_{i+1}\cdots P_{\mu-1}(I-\Pi_i)Q_{i+1} &= Q_iQ_{i+1}. \end{aligned}$$

Hence, multiplying (1.38) by $Q_iP_{i+1}\cdots P_{\mu-1}$, $i=0,\ldots,\mu-2$, and $Q_{\mu-1}$ yields

$$\begin{aligned} &Q_iP_{i+1}\cdots P_{\mu-1}G_\mu^{-1}B_\mu x(t) + Q_ix(t) - Q_iQ_{i+1}(\Pi_iQ_{i+1}x(t))' \\ &\quad - \sum_{l=i+1}^{\mu-2} Q_iP_{i+1}\cdots P_lQ_{l+1}(\Pi_lQ_{l+1}x(t))' = Q_iP_{i+1}\cdots P_{\mu-1}G_\mu^{-1}q(t) \end{aligned} \tag{1.40}$$

for $i=0,\ldots,\mu-2$ and

$$Q_{\mu-1}G_\mu^{-1}B_\mu x(t) + Q_{\mu-1}x(t) = Q_{\mu-1}G_\mu^{-1}q(t). \tag{1.41}$$

Equation (1.41) uniquely determines the component $Q_{\mu-1}x(t)$ as

$$Q_{\mu-1}x(t) = Q_{\mu-1}G_\mu^{-1}q(t) - Q_{\mu-1}G_\mu^{-1}B_\mu x(t),$$

and the formula contained in (1.40) for $i=\mu-2$ gives

$$\begin{aligned} &Q_{\mu-2}x(t) = \\ &\quad Q_{\mu-2}P_{\mu-1}G_\mu^{-1}q(t) - Q_{\mu-2}P_{\mu-1}G_\mu^{-1}B_\mu x(t) - Q_{\mu-2}Q_{\mu-1}(\Pi_{\mu-2}Q_{\mu-1}x(t))', \end{aligned}$$

and so on, i.e., in a consecutive manner we obtain expressions determining the components $Q_ix(t)$ with their dependence on $\Pi_{\mu-1}x(t)$ and $Q_{i+j}x(t)$, $j=1,\ldots,\mu-1-i$.

To compose an expression for the whole solution $x(t)$ there is no need for the components $Q_ix(t)$ themselves, $i=0,\ldots,\mu-1$. But one can do it with $Q_0x(t)$, $\Pi_{i-1}Q_ix(t)$, $i=1,\ldots,\mu-1$, which corresponds to the decomposition

$$I = Q_0 + \Pi_0Q_1 + \cdots + \Pi_{\mu-2}Q_{\mu-1} + \Pi_{\mu-1}. \tag{1.42}$$

For this purpose we rearrange the system (1.40), (1.41) once again by multiplying (1.41) by $\Pi_{\mu-2}$ and (1.40) for $i=1,\ldots,\mu-2$ by Π_{i-1}. Let us remark that, even though we scale with projectors (which are singular matrices) here, nothing of the equations gets lost. This is due to the relations

$$\begin{aligned} Q_i &= Q_i\Pi_{i-1}Q_i = (\Pi_{i-1} + (I - \Pi_{i-1}))Q_i\Pi_{i-1}Q_i \\ &\qquad\qquad\quad = (I + (I - \Pi_{i-1})Q_i)\Pi_{i-1}Q_i, \end{aligned} \tag{1.43}$$

$$\Pi_{i-1}Q_i = (I - (I - \Pi_{i-1})Q_i)Q_i,$$

which allow a one-to-one translation of the components $Q_i x(t)$ and $\Pi_{i-1}Q_i x(t)$ into each other. Choosing notation according to the decomposition (1.42),

$$v_0(t) := Q_0 x(t),\ \ v_i(t) := \Pi_{i-1}Q_i x(t),\ \ i = 1, \dots, \mu - 1,\ \ u(t) := \Pi_{i-1}x(t), \tag{1.44}$$

we obtain the representation, respectively decomposition

$$x(t) = v_0(t) + v_1(t) + \cdots + v_{\mu-1}(t) + u(t) \tag{1.45}$$

of the solution as well as the structured system resulting from (1.37), (1.40), and (1.41):

$$\begin{aligned} &\left[\begin{array}{c|cccc} I & & & & \\ \hline & 0 & \mathcal{N}_{01} & \cdots & \mathcal{N}_{0,\mu-1} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & \mathcal{N}_{\mu-2,\mu-1} \\ & & & & 0 \end{array}\right] \left[\begin{array}{c} u'(t) \\ \hline 0 \\ v_1'(t) \\ \vdots \\ v_{\mu-1}'(t) \end{array}\right] \\ &\qquad + \left[\begin{array}{c|ccc} \mathcal{W} & & & \\ \hline \mathcal{H}_0 & I & & \\ \vdots & & \ddots & \\ \vdots & & & \ddots \\ \mathcal{H}_{\mu-1} & & & I \end{array}\right] \left[\begin{array}{c} u(t) \\ \hline v_0(t) \\ \vdots \\ \vdots \\ v_{\mu-1}(t) \end{array}\right] = \left[\begin{array}{c} \mathcal{L}_d \\ \mathcal{L}_0 \\ \vdots \\ \vdots \\ \mathcal{L}_{\mu-1} \end{array}\right] q(t) \end{aligned} \tag{1.46}$$

with the $m \times m$ blocks

$$\begin{aligned} \mathcal{N}_{01} &:= -Q_0Q_1, & & \\ \mathcal{N}_{0j} &:= Q_0P_1 \cdots P_{j-1}Q_j, & & j = 2, \dots, \mu - 1, \\ \mathcal{N}_{i,i+1} &:= -\Pi_{i-1}Q_iQ_{i+1}, & & i = 1, \dots, \mu - 2, \\ \mathcal{N}_{ij} &:= -\Pi_{i-1}Q_iP_{i+1} \cdots P_{j-1}Q_j, & & j = i+2, \dots, \mu - 1,\ i = 1, \dots, \mu - 2, \end{aligned}$$

$$\begin{aligned} \mathcal{W} &:= \Pi_{\mu-1}G_\mu^{-1}B_\mu, & & \\ \mathcal{H}_0 &:= Q_0P_1 \cdots P_{\mu-1}G_\mu^{-1}B_\mu, & & \\ \mathcal{H}_i &:= \Pi_{i-1}Q_iP_{i+1} \cdots P_{\mu-1}G_\mu^{-1}B_\mu, & & i = 1, \dots, \mu - 2, \\ \mathcal{H}_{\mu-1} &:= \Pi_{\mu-2}Q_{\mu-1}G_\mu^{-1}B_\mu, & & \end{aligned}$$

and

$$\begin{aligned}
\mathcal{L}_d &:= \Pi_{\mu-1} G_\mu^{-1}, \\
\mathcal{L}_0 &:= Q_0 P_1 \cdots P_{\mu-1} G_\mu^{-1}, \\
\mathcal{L}_i &:= \Pi_{i-1} Q_i P_{i+1} \cdots P_{\mu-1} G_\mu^{-1}, \qquad i = 1, \ldots, \mu-2, \\
\mathcal{L}_{\mu-1} &:= \Pi_{\mu-2} Q_{\mu-1} G_\mu^{-1}.
\end{aligned}$$

System (1.46) almost looks like a DAE in Weierstraß–Kronecker form. However, compared to the latter it is a puffed up system of dimension $(\mu+1)m$. The system (1.46) is equivalent to the original DAE (1.30) in the following sense.

Proposition 1.20. *Let the DAE (1.30), with coefficients $E, F \in L(\mathbb{R}^m)$, have the characteristic values*

$$r_0 \leq \cdots \leq r_{\mu-1} < r_\mu = m.$$

(1) *If $x(.)$ is a solution of the DAE (1.30), then the components $u(.), v_0(.), \ldots, v_{\mu-1}(.)$ given by (1.44) form a solution of the puffed up system (1.46).*
(2) *Conversely, if the functions $u(.), v_0(.), \ldots, v_{\mu-1}(.)$ are a solution of the system (1.46) and if, additionally, $u(t_0) = \Pi_{\mu-1} u(t_0)$ holds for a $t_0 \in \mathcal{I}$, then the compound function $x(.)$ defined by (1.45) is a solution of the original DAE (1.30).*

Proof. It remains to verify (2). Due to the properties of the coefficients, for each solution of system (1.46) it holds that $v_i(t) = \Pi_{i-1} Q_i v_i(t)$, $i = 1, \ldots, \mu-1$, $v_0(t) = Q_0 v_0(t)$, which means that the components $v_i(t)$, $i = 0, \ldots, \mu-1$, belong to the desired subspaces.

The first equation in (1.46) is the explicit ODE $u'(t) + \mathcal{W} u(t) = \mathcal{L}_d q(t)$. Let $u_q(.)$ denote the solution fixed by the initial condition $u_q(t_0) = 0$. We have $u_q(t) = \Pi_{\mu-1} u_q(t)$ because of $\mathcal{W} = \Pi_{\mu-1} \mathcal{W}$, $\mathcal{L}_d = \Pi_{\mu-1} \mathcal{L}_d$. However, for each arbitrary constant $c \in \operatorname{im}(I - \Pi_{\mu-1})$, the function $\bar{u}(.) := c + u_q(.)$ solves this ODE but does not belong to $\operatorname{im} \Pi_{\mu-1}$ as we want it to.

With the initial condition $u(t_0) = u_0 \in \operatorname{im} \Pi_{\mu-1}$ the solution can be kept in the desired subspace, which means that $u(t) \in \operatorname{im} \Pi_{\mu-1}$ for all $t \in \mathcal{I}$. Now, by carrying out the decoupling procedure in reverse order and putting things together we have finished the proof. □

System (1.46) is *given in terms of the original DAE*. It shows in some detail the inherent structure of that DAE. It also serves as the idea of an analogous decoupling of time-varying linear DAEs (see Section 2.6).

Example 1.21 (Decoupling of an index-2 DAE). We reconsider the regular index-2 DAE

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} x' + \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} x = q$$

from Example 1.8, with the projectors

$$\Pi_1 = P_0P_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \; Q_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; P_0Q_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

The DAE itself can be rewritten without any differentiations of equations as

$$(-x_1 + x_3)' = q_2 + q_3 - q_1, \tag{1.47}$$

$$x_1' + x_2 = (q_1 - q_3), \tag{1.48}$$

$$x_1 + \frac{1}{2}(-x_1 + x_3) = \frac{1}{2}q_3. \tag{1.49}$$

Obviously, $\Pi_1 x$ reflects the proper state variable $-x_1 + x_3$, for which an explicit ODE (1.47) is given. P_0Q_1x refers to the variable x_1 that is described by the algebraic equation (1.49) when the solution $-x_1 + x_3$ is already given by (1.47). Finally, Q_0x reflects the variable x_2 which can be determined by (1.48). Note, that the variable x_1 has to be differentiated here. Simple calculations yield $\mathcal{W} = \Pi_{\mu-1}G_2^{-1}B_0\Pi_{\mu-1} = 0$, $\mathcal{H}_0 = Q_0P_1G_2^{-1}B_0\Pi_{\mu-1} = 0$ and

$$\mathcal{H}_1 = Q_1G_2^{-1}B_0\Pi_{\mu-1} = \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}.$$

This way the DAE decouples as

$$(\Pi_1 x)' = \Pi_1 G_2^{-1} q, \tag{1.50}$$

$$-Q_0Q_1(\Pi_0Q_1x)' + Q_0x = Q_0P_1G_2^{-1}q, \tag{1.51}$$

$$\Pi_0Q_1 + \mathcal{H}_1\Pi_1 x = \Pi_0Q_1G_2^{-1}q. \tag{1.52}$$

These equations mean in full detail

$$\left(\begin{bmatrix} 0 \\ 0 \\ -x_1 + x_3 \end{bmatrix}\right)' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 1 & 1 \end{bmatrix} q,$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\left(\begin{bmatrix} x_1 \\ 0 \\ x_1 \end{bmatrix}\right)' + \begin{bmatrix} 0 \\ x_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} q,$$

$$\begin{bmatrix} x_1 \\ 0 \\ x_1 \end{bmatrix} + \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ -x_1 + x_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} q.$$

Dropping the redundant equations as well as all zero lines one arrives exactly at the compressed form (1.47)–(1.49). □

1.2.3 Complete decoupling

A special smart choice of the admissible projectors cancels the coefficients $\mathcal{H}_i$ in system (1.46) so that the second part no longer depends on the first part.

Theorem 1.22. *Let $\{E,F\}$, $E,F \in L(\mathbb{R}^m)$, be a pair with characteristic values*

$$r_0 \leq \cdots \leq r_{\mu-1} < r_\mu = m.$$

Then there are admissible projectors $Q_0, \ldots Q_{\mu-1}$ such that the coupling coefficients $\mathcal{H}_0, \ldots, \mathcal{H}_{\mu-1}$ in (1.46) vanish, that is, (1.46) decouples into two independent subsystems.

Proof. For any given sequence of admissible projectors $Q_0, \ldots, Q_{\mu-1}$ the coupling coefficients can be expressed as $\mathcal{H}_0 = Q_{0*}\Pi_{\mu-1}$ and $\mathcal{H}_i = \Pi_{i-1}Q_{i*}\Pi_{\mu-1}$ for $i = 1, \ldots, \mu-1$, where we denote

$$\begin{aligned} Q_{0*} &:= Q_0 P_1 \cdots P_{\mu-1} G_\mu^{-1} B_0, \\ Q_{i*} &:= Q_i P_{i+1} \cdots P_{\mu-1} G_\mu^{-1} B_0 \Pi_{i-1}, \qquad i = 1, \ldots, \mu-2, \\ Q_{\mu-1*} &:= Q_{\mu-1} G_\mu^{-1} B_0 \Pi_{\mu-2}. \end{aligned}$$

We realize that $Q_{i*}Q_i = Q_i$, $i = 0, \ldots, \mu-1$, since

$$\begin{aligned} Q_{\mu-1*}Q_{\mu-1} = Q_{\mu-1}G_\mu^{-1}B_0\Pi_{\mu-2}Q_{\mu-1} &= Q_{\mu-1}G_\mu^{-1}B_{\mu-1}Q_{\mu-1} \\ &= Q_{\mu-1}G_\mu^{-1}G_\mu Q_{\mu-1} = Q_{\mu-1}, \end{aligned}$$

and so on for $i = \mu-2, \ldots, 0$. This implies $(Q_{i*})^2 = Q_{i*}$, i.e., Q_{i*} is a projector onto N_i, $i = 0, \ldots, \mu-1$. By construction one has $N_0 + \cdots + N_{i-1} \subseteq \ker Q_{i*}$ for $i = 1, \ldots, \mu-1$. The new projectors $\bar{Q}_0 := Q_0, \ldots, \bar{Q}_{\mu-2} := Q_{\mu-2}$, $\bar{Q}_{\mu-1} := Q_{\mu-1*}$ are also admissible, but now, the respective coefficient $\bar{\mathcal{H}}_{\mu-1}$ disappears in (1.46). Namely, the old and new sequences are related by

$$\bar{G}_i = G_i, \;\; i = 0, \ldots, \mu-1, \;\; \bar{G}_\mu = G_\mu + B_{\mu-1}Q_{\mu-1*} = G_\mu Z_\mu$$

with nonsingular $Z_\mu := I + Q_{\mu-1}Q_{\mu-1*}P_{\mu-1}$. This yields

$$\begin{aligned} \bar{Q}_{\mu-1*} := \bar{Q}_{\mu-1}\bar{G}_{\mu-1}B_0\Pi_{\mu-2} &= Q_{\mu-1*}Z_\mu^{-1}G_\mu^{-1}B_0\Pi_{\mu-2} \\ &= Q_{\mu-1}G_\mu^{-1}B_0\Pi_{\mu-2} = Q_{\mu-1*} = \bar{Q}_{\mu-1} \end{aligned}$$

because of

$$Q_{\mu-1*}Z_\mu^{-1} = Q_{\mu-1*}(I - Q_{\mu-1}Q_{\mu-1*}P_{\mu-1}) = Q_{\mu-1},$$

and hence

$$\bar{\mathcal{H}}_{\mu-1} := \bar{\Pi}_{\mu-2}\bar{Q}_{\mu-1*}\bar{\Pi}_{\mu-1} = \Pi_{\mu-2}\bar{Q}_{\mu-1}\bar{\Pi}_{\mu-1} = 0.$$

We show by induction that the coupling coefficients disappear stepwise with an appropriate choice of admissible projectors. Assume $Q_0, \ldots, Q_{\mu-1}$ to be such that

$$\mathcal{H}_{k+1} = 0, \dots, \mathcal{H}_{\mu-1} = 0, \tag{1.53}$$

or, equivalently,

$$Q_{k+1*}\Pi_{\mu-1} = 0, \dots, Q_{\mu-1*}\Pi_{\mu-1} = 0,$$

for a certain k, $0 \le k \le \mu - 2$. We build a new sequence by letting $\bar{Q}_i := Q_i$ for $i = 0, \dots, k-1$ (if $k \ge 1$) and $\bar{Q}_k := Q_{k*}$. Thus, $Q_k\bar{P}_k = -\bar{Q}_k P_k$ and the projectors $\bar{Q}_0, \dots, \bar{Q}_k$ are admissible. The resulting two sequences are related by

$$\bar{G}_i = G_i Z_i, \qquad i = 0, \dots, k+1,$$

with factors

$$Z_0 = I, \quad \dots, \quad Z_k = I, \quad Z_{k+1} = I + Q_k Q_{k*} P_k, \quad Z_{k+1}^{-1} = I - Q_k Q_{k*} P_k.$$

We form $\bar{Q}_{k+1} := Z_{k+1}^{-1} Q_{k+1} Z_{k+1} = Z_{k+1}^{-1} Q_{k+1}$. Then, $\bar{Q}_0, \dots, \bar{Q}_{k+1}$ are also admissible. Applying Lemma 1.18 we proceed with

$$\bar{G}_j = G_j Z_j, \quad \bar{Q}_j := Z_j^{-1} Q_j Z_j, \quad j = k+2, \dots, \mu - 1,$$

and arrive at a new sequence of admissible projectors $\bar{Q}_0, \dots, \bar{Q}_{\mu-1}$. The invertibility of Z_j is ensured by Lemma 1.18. Putting $Y_{k+1} := Z_{k+1}$ and, exploiting Lemma 1.18,

$$Y_j := Z_j Z_{j-1}^{-1} = I + Q_{j-1}(\bar{\Pi}_{j-2}\bar{Q}_{j-1} - \Pi_{j-2}Q_{j-1}) + \sum_{l=0}^{j-2} Q_l \bar{\Pi}_{j-2}\bar{Q}_{j-1}, \quad j \ge k+2.$$

Additionally, we learn from Lemma 1.18 that the subspaces $N_0 \oplus \cdots \oplus N_j$ and $\bar{N}_0 \oplus \cdots \oplus \bar{N}_j$ coincide. The expression for $Y_j,\ j \ge k+2$, simplifies to

$$Y_j = I + \sum_{l=0}^{j-2} Q_l \bar{\Pi}_{j-2}\bar{Q}_{j-1} = I + \sum_{l=k}^{j-2} Q_l \bar{\Pi}_{j-2} Q_{j-1}$$

for our special new projectors because the following relations are valid:

$$Q_j Z_j = 0, \quad \bar{Q}_j = Z_j^{-1} Q_j, \quad \bar{\Pi}_{j-2}\bar{Q}_{j-1} = \bar{\Pi}_{j-2} Z_{j-1}^{-1} Q_{j-1} = \bar{\Pi}_{j-2} Q_{j-1},$$
$$Q_{j-1}(\bar{\Pi}_{j-2}\bar{Q}_{j-1} - \Pi_{j-2}Q_{j-1}) = Q_{j-1}(\bar{\Pi}_{j-2}Q_{j-1} - \Pi_{j-2}Q_{j-1}) = 0.$$

We have to verify that the new coupling coefficients $\bar{\mathcal{H}}_k$ and $\bar{\mathcal{H}}_j,\ j \ge k+1$, disappear. We compute $\bar{Q}_k Z_{k+1}^{-1} = \bar{Q}_k - \bar{Q}_k P_k = \bar{Q}_k Q_k = Q_k$ and

$$Z_{j-1} Z_j^{-1} = Y_j^{-1} = I - \sum_{l=k}^{j-2} Q_l \bar{\Pi}_{j-2} Q_{j-1}, \quad j \ge k+2. \tag{1.54}$$

For $j \ge k+1$ this yields

$$\bar{Q}_{j*}\bar{\Pi}_{\mu-1} = \bar{Q}_j \bar{P}_{j+1} \cdots \bar{P}_{\mu-1} \bar{G}_\mu^{-1} B \bar{\Pi}_{\mu-1} = Z_j^{-1} Q_j Y_{j+1}^{-1} P_{j+1} \cdots Y_{\mu-1}^{-1} P_{\mu-1} Y_\mu^{-1} B \bar{\Pi}_{\mu-1}$$

and, by inserting (1.54) into the last expression,

$$\bar{Q}_{j*}\bar{\Pi}_{\mu-1} =$$
$$Z_j^{-1}Q_j(I-\sum_{l=k}^{j-1}Q_l\bar{\Pi}_{j-1}Q_j)P_{j+1}\cdots P_{\mu-1}(I-\sum_{l=k}^{\mu-2}Q_l\bar{\Pi}_{\mu-2}Q_{\mu-1})G_\mu^{-1}B\bar{\Pi}_{\mu-1}.$$

Rearranging the terms one finds

$$\begin{aligned}\bar{Q}_{j*}\bar{\Pi}_{\mu-1} = (Z_j^{-1}Q_jP_{j+1}\cdots P_{\mu-1} + \mathcal{C}_{j,j+1}Q_{j+1}P_{j+2}\cdots P_{\mu-1}\\ +\cdots+\mathcal{C}_{j,\mu-2}Q_{\mu-2}P_{\mu-1}+\mathcal{C}_{j,\mu-1}Q_{\mu-1})G_\mu^{-1}B\bar{\Pi}_{\mu-1}.\end{aligned} \tag{1.55}$$

The detailed expression of the coefficients $\mathcal{C}_{j,i}$ does not matter at all. With analogous arguments we derive

$$\begin{aligned}\bar{Q}_{k*}\bar{\Pi}_{\mu-1} = (Q_{k*}P_{k+1}\cdots P_{\mu-1} + \mathcal{C}_{k,j+1}Q_{k+1}P_{k+2}\cdots P_{\mu-1}\\ +\cdots+\mathcal{C}_{k,\mu-2}Q_{\mu-2}P_{\mu-1}+\mathcal{C}_{k,\mu-1}Q_{\mu-1})G_\mu^{-1}B\bar{\Pi}_{\mu-1}.\end{aligned} \tag{1.56}$$

Next we compute

$$\begin{aligned}\bar{\Pi}_{\mu-1} &= \Pi_{k-1}\bar{P}_k\bar{P}_{k+1}\cdots\bar{P}_{\mu-1} = \Pi_{k-1}\bar{P}_kP_{k+1}\cdots P_{\mu-1}\\ &= \Pi_{k-1}(P_k+Q_k)\bar{P}_kP_{k+1}\cdots P_{\mu-1} = \Pi_{\mu-1} - Q_k\bar{Q}_k\Pi_{\mu-1},\end{aligned}$$

and therefore

$$G_\mu^{-1}B\bar{\Pi}_{\mu-1} = G_\mu^{-1}B(\Pi_{\mu-1}-\Pi_{k-1}Q_k\bar{Q}_k\Pi_{\mu-1}) = G_\mu^{-1}B\Pi_{\mu-1} - Q_k\bar{Q}_k\Pi_{\mu-1}.$$

Regarding assumption (1.53) and the properties of admissible projectors we have

$$Q_{\mu-1}G_\mu^{-1}B\bar{\Pi}_{\mu-1} = Q_{\mu-1}G_\mu^{-1}B\Pi_{\mu-1} - Q_{\mu-1}\bar{Q}_k\Pi_{\mu-1} = Q_{\mu-1*}\Pi_{\mu-1} = 0,$$

and, for $i = k+1,\ldots,\mu-2$,

$$Q_iP_{i+1}\cdots P_{\mu-1}B\bar{\Pi}_{\mu-1} = Q_iP_{i+1}\cdots P_{\mu-1}B\Pi_{\mu-1} - Q_i\bar{Q}_k\Pi_{\mu-1} = Q_{i*}\Pi_{\mu-1} = 0.$$

Furthermore, taking into account the special choice of $\bar{Q}_k$,

$$\begin{aligned}Q_kP_{k+1}\cdots P_{\mu-1}B\bar{\Pi}_{\mu-1} &= Q_kP_{k+1}\cdots P_{\mu-1}B\Pi_{\mu-1} - Q_k\bar{Q}_k\Pi_{\mu-1}\\ &= (Q_{k*}-\bar{Q}_k)\Pi_{\mu-1} = 0.\end{aligned}$$

This makes it evident that all single summands on the right-hand sides of the formulas (1.55) and (1.56) disappear, and thus $\bar{Q}_{j*}\bar{\Pi}_{\mu-1} = 0$ for $j = k,\ldots,\mu-1$, that is, the new decoupling coefficients vanish. In consequence, starting with any admissible projectors we apply the above procedure first for $k = \mu-1$, then for $k = \mu-2$ up to $k = 0$. At each level an additional coupling coefficient is canceled, and we finish with a complete decoupling of the two parts in (1.46). □

Definition 1.23. Let the DAE (1.30), with coefficients $E,F \in L(\mathbb{R}^m)$, have the structural characteristic values

$$r_0 \leq \cdots \leq r_{\mu-1} < r_\mu = m,$$

and let the system (1.46) be generated by an admissible matrix sequence $G_0,\ldots,G_\mu$. If in (1.46) all coefficients $\mathcal{H}_i$, $i=0,\ldots,\mu-1$, vanish, then the underlying admissible projectors $Q_0,\ldots,Q_{\mu-1}$ are called *completely decoupling projectors* for the DAE (1.30).

The completely decoupled system (1.46) offers as much insight as the Weierstraß-Kronecker form does.

Example 1.24 (Complete decoupling of an index-2 DAE). We reconsider once more the regular index-2 DAE

$$\begin{bmatrix}1&0&0\\0&0&1\\0&0&0\end{bmatrix}x' + \begin{bmatrix}1&1&1\\0&1&0\\1&0&1\end{bmatrix}x = q$$

from Examples 1.8 and 1.21. The previously used projectors do not yield a complete decoupling. We now use a different projector Q_1 such that

$$Q_1 = \begin{bmatrix}\frac{1}{2}&0&\frac{1}{2}\\-\frac{1}{2}&0&-\frac{1}{2}\\\frac{1}{2}&0&\frac{1}{2}\end{bmatrix}, \quad G_2 = \begin{bmatrix}2&1&1\\0&1&1\\1&0&1\end{bmatrix},$$

and further

$$\Pi_1 = P_0P_1 = \begin{bmatrix}\frac{1}{2}&0&-\frac{1}{2}\\0&0&0\\-\frac{1}{2}&0&\frac{1}{2}\end{bmatrix}, Q_0 = \begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix}, P_0Q_1 = \begin{bmatrix}\frac{1}{2}&0&\frac{1}{2}\\0&0&0\\\frac{1}{2}&0&\frac{1}{2}\end{bmatrix}.$$

The DAE itself can be rewritten without any differentiations of equations as

$$\begin{aligned}(x_1-x_3)' &= q_1-q_2-q_3,\\(x_1+x_3)'+2x_2 &= q1+q2-q_3,\\x_1+x_3 &= q_3.\end{aligned}$$

Obviously, $\Pi_1 x$ again reflects the proper state variable $-x_1+x_3$, for which an explicit ODE is given. P_0Q_1x refers to the variable x_1+x_3 that is described by the algebraic equation. Finally, Q_0x reflects the variable x_2. Simple calculations yield $\mathcal{W} = \Pi_{\mu-1}G_2^{-1}B_0\Pi_{\mu-1} = 0$, $\mathcal{H}_0 = Q_0P_1G_2^{-1}B_0\Pi_{\mu-1} = 0$ and $\mathcal{H}_1 = Q_1G_2^{-1}B_0\Pi_{\mu-1} = 0$. In this way the DAE decouples completely as

$$\begin{aligned}(\Pi_1 x)' &= \Pi_1 G_2^{-1} q,\\ -Q_0 Q_1 (\Pi_0 Q_1 x)' + Q_0 x &= Q_0 P_1 G_2^{-1} q,\\ \Pi_0 Q_1 &= \Pi_0 Q_1 G_2^{-1} q.\end{aligned}$$

These equations mean in full detail

$$\left(\begin{bmatrix}\frac{1}{2}(x_1-x_3)\\ 0\\ -\frac{1}{2}(x_1-x_3)\end{bmatrix}\right)' = \begin{bmatrix}\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2}\\ 0 & 0 & 0\\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2}\end{bmatrix} q,$$

$$\begin{bmatrix}0&0&0\\ \frac{1}{2}&0&\frac{1}{2}\\ 0&0&0\end{bmatrix}\left(\begin{bmatrix}\frac{1}{2}(x_1+x_3)\\ 0\\ \frac{1}{2}(x_1+x_3)\end{bmatrix}\right)' + \begin{bmatrix}0\\ x_2\\ 0\end{bmatrix} = \begin{bmatrix}0&0&0\\ \frac{1}{2}&\frac{1}{2}&-\frac{1}{2}\\ 0&0&0\end{bmatrix} q,$$

$$\begin{bmatrix}\frac{1}{2}(x_1+x_3)\\ 0\\ \frac{1}{2}(x_1+x_3)\end{bmatrix} = \begin{bmatrix}0&0&\frac{1}{2}\\ 0&0&0\\ 0&0&\frac{1}{2}\end{bmatrix} q.$$

Dropping the redundant equations as well as all zero lines one arrives exactly at the compressed form described above. □

Example 1.25 (Decoupling of the DAE in Example 1.5). The following matrix sequence is admissible for the pair $\{E, F\}$ from Example 1.5 which is regular with index 4:

$$G_0 = E = \begin{bmatrix}1&0&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\\0&0&0&0&0\end{bmatrix},\quad Q_0 = \begin{bmatrix}0&0&0&0&0\\0&1&0&0&0\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix},\quad B_0 = F = \begin{bmatrix}-\alpha&-1&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix},$$

$$G_1 = \begin{bmatrix}1&-1&0&0&0\\0&1&1&0&0\\0&0&0&1&0\\0&0&0&0&1\\0&0&0&0&0\end{bmatrix},\quad Q_1 = \begin{bmatrix}0&0&-1&0&0\\0&0&-1&0&0\\0&0&1&0&0\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix},\quad \Pi_0 Q_1 = \begin{bmatrix}0&0&-1&0&0\\0&0&0&0&0\\0&0&1&0&0\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix},$$

$$G_2 = \begin{bmatrix}1&-1&\alpha&0&0\\0&1&1&0&0\\0&0&1&1&0\\0&0&0&0&1\\0&0&0&0&0\end{bmatrix},\quad Q_2 = \begin{bmatrix}0&0&0&1+\alpha&0\\0&0&0&1&0\\0&0&0&-1&0\\0&0&0&1&0\\0&0&0&0&0\end{bmatrix},\quad \Pi_1 Q_2 = \begin{bmatrix}0&0&0&\alpha&0\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&1&0\\0&0&0&0&0\end{bmatrix},$$

$$G_3 = \begin{bmatrix} 1 & -1 & \alpha & -\alpha^2 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \; Q_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & -1-\alpha-\alpha^2 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \; \Pi_2 Q_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & -\alpha^2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$G_4 = \begin{bmatrix} 1 & -1 & \alpha & -\alpha^2 & \alpha^3 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \Pi_3 = \begin{bmatrix} 1 & 0 & 1 & -\alpha & -\alpha^2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

and the characteristic values are $r_0 = r_1 = r_2 = r_3 = 4, r_4 = 5$ and $\mu = 4$. Additionally, it follows that

$$\begin{aligned} Q_3 G_4^{-1} B_0 \Pi_3 &= 0, & Q_2 P_3 G_4^{-1} B_0 \Pi_3 &= 0, \\ Q_1 P_2 P_3 G_4^{-1} B_0 \Pi_3 &= 0, & Q_0 P_1 P_2 P_3 G_4^{-1} B_0 \Pi_3 &= 0, \end{aligned}$$

and

$$\Pi_3 G_4^{-1} B_0 \Pi_3 = -\alpha \Pi_3. \tag{1.57}$$

The projectors Q_0, Q_1, Q_2, Q_3 provide a complete decoupling of the given DAE $Ex'(t) + Fx(t) = q(t)$. The projectors Q_0, $\Pi_0 Q_1$, $\Pi_1 Q_2$ and $\Pi_2 Q_3$ represent the variables x_2, x_3, x_4 and x_5, respectively. The projector Π_3 and the coefficient (1.57) determine the inherent regular ODE, namely (the zero rows are dropped)

$$(x_1 + x_3 - \alpha x_4 + \alpha^2 x_5)' - \alpha(x_1 + x_3 - \alpha x_4 + \alpha^2 x_5) = q_1 + q_2 - \alpha q_3 + \alpha^2 q_4 - \alpha^3 q_5.$$

It is noteworthy that no derivatives of the excitation q encroach in this ODE. □

Notice that for DAEs with $\mu = 1$, the completely decoupling projector Q_0 is uniquely determined. It is the projector onto N_0 along $S_0 = \{z \in \mathbb{R}^m : B_0 z \in \operatorname{im} G_0\}$ (cf. Appendix A). However, for higher index $\mu > 1$, there are many complete decouplings, as the next example shows.

Example 1.26 (Diversity of completely decoupling projectors). Let

$$E = G_0 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad F = B_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and choose projectors with a free parameter α:

$$Q_0 = \begin{bmatrix} 1 & \alpha & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad P_0 = \begin{bmatrix} 1 & -\alpha & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad G_1 = \begin{bmatrix} 1 & 1+\alpha & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; B_1 = P_0,$$

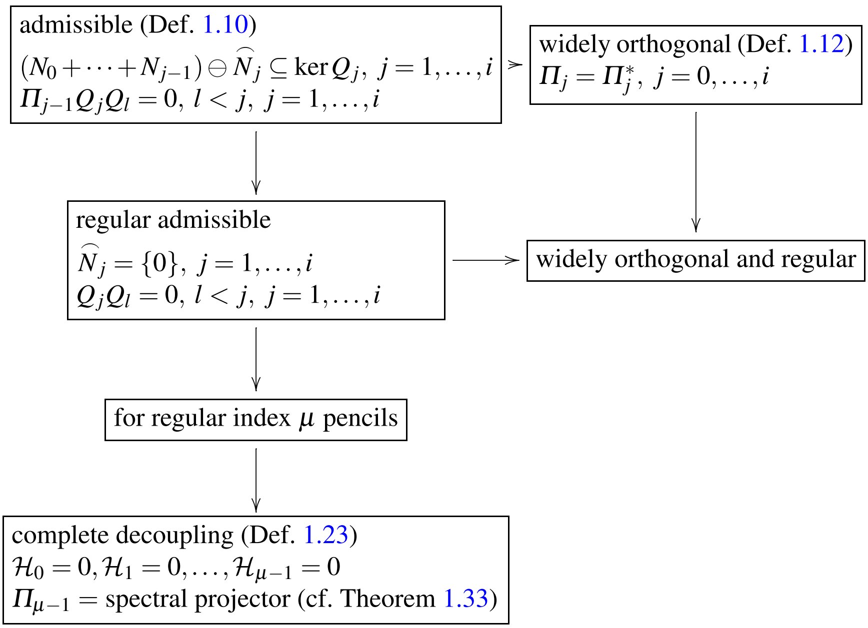
$$Q_1 = \begin{bmatrix} 0 & -(1+\alpha) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \Pi_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad G_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$G_2^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad Q_0 P_1 G_2^{-1} B_0 = Q_0,$$

i.e., Q_0 and Q_1 are completely decoupling projectors for each arbitrary value α. However, in contrast, the projector Π_1 is independent of α. □

1.2.4 Hierarchy of projector sequences for constant matrix pencils

The matrices $Q_0, \dots, Q_i$ are projectors, where Q_j projects onto $N_j = \ker G_j$, $j = 0, \dots, i$, with $P_0 := I - Q_0$, $\Pi_0 := P_0$ and $P_j := I - Q_j$, $\Pi_j := \Pi_{j-1} P_j$, $\widehat{N}_j := (N_0 + \cdots + N_{j-1}) \cap N_j$, $j = 1, \dots, i$.



1.2.5 Compression to a generalized Weierstraß–Kronecker form

The DAE (1.30) as well as its decoupled version (1.35) comprise m equations. The advanced decoupled system (1.46) is formally composed of $m(\mu+1)$ equations; however, it can be compressed back on an m-dimensional DAE without losing information. The next lemma records essential properties to be used in the compression procedure.

Lemma 1.27. *The entries $\mathcal{N}_{ij}$ of the decoupled system (1.46) have the following properties for $i = 0,\dots,\mu-2$:*

$$\begin{aligned}
\mathcal{N}_{i,i+1} &= \mathcal{N}_{i,i+1}\Pi_i Q_{i+1},\\
\mathcal{N}_{ij} &= \mathcal{N}_{ij}\Pi_{j-1}Q_j, \quad j = i+2,\dots,\mu-1,\\
\ker\mathcal{N}_{i,i+1} &= \ker\Pi_i Q_{i+1},\\
\operatorname{rank}\mathcal{N}_{i,i+1} &= m - r_{i+1}.
\end{aligned}$$

Proof. We use the additional subspaces $S_i := \ker\mathcal{W}_i B_i \subseteq \mathbb{R}^m$ and the projectors $\mathcal{W}_i \in L(\mathbb{R}^m)$ with

$$\ker\mathcal{W}_i = \operatorname{im} G_i,\ i = 0,\dots,\mu-1.$$

Let G_i^- be the generalized reflexive inverse of G_i with $G_iG_i^-G_i = G_i$, $G_i^-G_iG_i^- = G_i^-$, $G_iG_i^- = I - \mathcal{W}_i$ and $G_i^-G_i = P_i$. We factorize G_{i+1} as

$$G_{i+1} = G_i + B_iQ_i = G_i + \mathcal{W}_iB_iQ_i + G_iG_i^-B_iQ_i = \mathcal{G}_{i+1}\mathcal{F}_{i+1},$$

$$\mathcal{G}_{i+1} := G_i + \mathcal{W}_iB_iQ_i, \quad \mathcal{F}_{i+1} = I + P_iG_i^-B_iQ_i.$$

Since $\mathcal{F}_{i+1}$ is invertible (cf. Lemma A.3), it follows that $\mathcal{G}_{i+1}$ has rank r_{i+1} like G_{i+1}.

Furthermore, it holds that $\ker\mathcal{G}_{i+1} = N_i \cap S_i$. Namely, $\mathcal{G}_{i+1}z = 0$ means that $G_iz = 0$ and $\mathcal{W}_iB_iQ_iz = 0$, i.e., $z = Q_iz$ and $\mathcal{W}_iB_iz = 0$, but this is $z \in N_i \cap S_i$. Therefore, $N_i \cap S_i$ must have the dimension $m - r_{i+1}$. Next we derive the relation

$$N_i \cap S_i = \operatorname{im} Q_iQ_{i+1}. \tag{1.58}$$

If $z \in N_i \cap S_i$ then $z = Q_iz$ and $B_iz = G_iw$ implying $(G_i + B_iQ_i)(P_iw + Q_iz) = 0$, and hence $P_iw + Q_iz = Q_{i+1}(P_iw + Q_iz) = Q_{i+1}w$. Therefore, $z = Q_iz = Q_iQ_{i+1}w$. Consequently, $N_i \cap S_i \subseteq \operatorname{im} Q_iQ_{i+1}$. Conversely, assume $z = Q_iQ_{i+1}y$. Taking into consideration that $(G_i + B_iQ_i)Q_{i+1} = 0$, we derive $z = Q_iz$ and $B_iz = B_iQ_iQ_{i+1}y = -G_iQ_{i+1}y$, i.e., $z \in N_i$ and $z \in S_i$. Thus, relation (1.58) is valid.

Owing to (1.58) we have

$$\operatorname{rank} Q_iQ_{i+1} = \dim N_i \cap S_i = m - r_{i+1}. \tag{1.59}$$

If follows immediately that $\operatorname{rank}\mathcal{N}_{i,i+1} = m - r_{i+1}$, and, since $\operatorname{im} P_{i+1} \subseteq \ker\mathcal{N}_{i,i+1}$, $\operatorname{rank} P_{i+1} = r_{i+1}$, that $\operatorname{im} P_{i+1} = \ker\mathcal{N}_{i,i+1}$. □

We turn to the compression of the large system (1.46) on m dimensions. The projector Q_0 has rank $m-r_0$, the projector $\Pi_{i-1}Q_i$ has rank $m-r_i$ for $i=1,\dots,\mu-1$, and $\Pi_{\mu-1}$ has rank $d:=m-\sum_{j=0}^{\mu-1}(m-r_j)$.

We introduce full-row-rank matrices $\Gamma_i\in L(\mathbb{R}^m,\mathbb{R}^{m-r_i})$, $i=0,\dots,\mu-1$, and $\Gamma_d\in L(\mathbb{R}^m,\mathbb{R}^d)$ such that

$$\begin{aligned}
\operatorname{im}\Gamma_d\Pi_{\mu-1} &= \Gamma_d\operatorname{im}\Pi_{\mu-1}=\mathbb{R}^d, & \ker\Gamma_d &= \operatorname{im}(I-\Pi_{\mu-1})=N_0+\cdots+N_{\mu-1},\\
\Gamma_0N_0 &= \mathbb{R}^{m-r_0}, & \ker\Gamma_0 &= \ker Q_0,\\
\Gamma_i\Pi_{i-1}N_i &= \mathbb{R}^{m-r_i}, & \ker\Gamma_i &= \ker\Pi_{i-1}Q_i,\ i=1,\dots,\mu-1,
\end{aligned}$$

as well as generalized inverses Γ_d^-,Γ_i^-, $i=0,\dots,\mu-1$, such that

$$\begin{aligned}
\Gamma_d^-\Gamma_d &= \Pi_{\mu-1}, & \Gamma_d\Gamma_d^- &= I,\\
\Gamma_i^-\Gamma_i &= \Pi_{i-1}Q_i, & \Gamma_i\Gamma_i^- &= I, \qquad i=1,\dots,\mu-1,\\
\Gamma_0^-\Gamma_0 &= Q_0, & \Gamma_0\Gamma_0^- &= I.
\end{aligned}$$

If the projectors $Q_0,\dots,Q_{\mu-1}$ are widely orthogonal (cf. Proposition 1.13(6)), then the above projectors are symmetric and Γ_d^-, Γ_i^- are the Moore–Penrose generalized inverses. Denoting

$$\tilde{\mathcal{H}}_i:=\Gamma_i\mathcal{H}_i\Gamma_d^-,\qquad \tilde{\mathcal{L}}_i:=\Gamma_i\mathcal{L}_i,\ i=0,\dots,\mu-1, \tag{1.60}$$

$$\tilde{\mathcal{W}}:=\Gamma_d\mathcal{W}\Gamma_d^-,\qquad \tilde{\mathcal{L}}_d:=\Gamma_d\mathcal{L}_d, \tag{1.61}$$

$$\tilde{\mathcal{N}}_{ij}:=\Gamma_i\mathcal{N}_{ij}\Gamma_j^-,\ \ j=i+1,\dots,\mu-1,\ i=0,\dots,\mu-2, \tag{1.62}$$

and transforming the new variables

$$\tilde{u}=\Gamma_du,\qquad \tilde{v}_i=\Gamma_iv_i,\qquad i=0,\dots,\mu-1, \tag{1.63}$$

$$u=\Gamma_d^-\tilde{u},\qquad v_i=\Gamma_i^-\tilde{v}_i,\qquad i=0,\dots,\mu-1, \tag{1.64}$$

we compress the large system (1.46) into the m-dimensional one

$$\left[\begin{array}{c|cccc}
I & & & & \\ \hline
 & 0 & \tilde{\mathcal{N}}_{01} & \cdots & \tilde{\mathcal{N}}_{0,\mu-1}\\
 & & \ddots & \ddots & \vdots\\
 & & & \ddots & \tilde{\mathcal{N}}_{\mu-2,\mu-1}\\
 & & & & 0
\end{array}\right]
\left[\begin{array}{c}\tilde{u}'(t)\\ \hline 0\\ \tilde{v}_1'(t)\\ \vdots\\ \tilde{v}_{\mu-1}'(t)\end{array}\right]
+\left[\begin{array}{c|ccc}
\tilde{\mathcal{W}} & & & \\ \hline
\tilde{\mathcal{H}}_0 & I & & \\
\vdots & & \ddots & \\
\vdots & & & \ddots \\
\tilde{\mathcal{H}}_{\mu-1} & & & I
\end{array}\right]
\left[\begin{array}{c}\tilde{u}(t)\\ \hline \tilde{v}_0(t)\\ \vdots\\ \vdots\\ \tilde{v}_{\mu-1}(t)\end{array}\right]
=\left[\begin{array}{c}\tilde{\mathcal{L}}_d\\ \tilde{\mathcal{L}}_0\\ \vdots\\ \vdots\\ \tilde{\mathcal{L}}_{\mu-1}\end{array}\right]q \tag{1.65}$$

without losing any information. As a consequence of Lemma 1.27, the blocks $\tilde{\mathcal{N}}_{i,i+1}$ have full column rank $m - r_{i+1}$ for $i = 0,\dots,\mu-2$.

Proposition 1.28. *Let the pair $\{E,F\}$, $E,F \in L(\mathbb{R}^m)$ have the structural characteristic values*

$$r_0 \le \cdots \le r_{\mu-1} < r_\mu = m.$$

(1) *Then there are nonsingular matrices $L,K \in L(\mathbb{R}^m)$ such that*

$$LEK = \left[\begin{array}{c|cccc} I & & & & \\ \hline & 0 & \tilde{\mathcal{N}}_{01} & \cdots & \tilde{\mathcal{N}}_{0,\mu-1} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & \tilde{\mathcal{N}}_{\mu-2,\mu-1} \\ & & & & 0 \end{array}\right], \quad LFK = \left[\begin{array}{c|ccc} \tilde{\mathcal{W}} & & & \\ \hline \tilde{\mathcal{H}}_0 & I & & \\ \vdots & & \ddots & \\ \vdots & & & \ddots \\ \tilde{\mathcal{H}}_{\mu-1} & & & I \end{array}\right],$$

with entries described by (1.60)–(1.62). Each block $\tilde{\mathcal{N}}_{i,i+1}$ has full column rank $m - r_{i+1}$, $i = 0,\dots,\mu-2$, and hence the nilpotent part in LEK has index μ.

(2) *By means of completely decoupling projectors, L and K can be built so that the coefficients $\tilde{\mathcal{H}}_0,\dots,\tilde{\mathcal{H}}_{\mu-1}$ disappear, and the DAE transforms into Weierstraß–Kronecker form (1.3) with $l = \sum_{i=0}^{\mu-1}(m - r_i)$.*

Proof. Due to the properties

$$\begin{aligned} \mathcal{H}_i &= \mathcal{H}_i\Pi_{\mu-1} = \mathcal{H}_i\Gamma_d^-\Gamma_d, \quad i = 0,\dots,\mu-1, \\ \mathcal{W} &= \mathcal{W}\Pi_{\mu-1} = \mathcal{W}\Gamma_d^-\Gamma_d, \\ \mathcal{N}_{ij} &= \mathcal{N}_{ij}\Pi_{j-1}Q_j = \mathcal{N}_{ij}\Gamma_j^-\Gamma_j, \quad j = 1,\dots,\mu-1, \ i = 0,\dots,\mu-2, \end{aligned}$$

we can recover system (1.46) from (1.65) by multiplying on the left by

$$\Gamma^- := \left[\begin{array}{c|ccc} \Gamma_d^- & & & \\ \hline & \Gamma_0^- & & \\ & & \ddots & \\ & & & \Gamma_{\mu-1}^- \end{array}\right] \in L(\mathbb{R}^m, \mathbb{R}^{(\mu+1)m})$$

using transformation (1.64) and taking into account that $u = \Gamma_d^-\tilde{u} = \Pi_{\mu-1}u$ and $\Pi_{\mu-1}u' = u'$. The matrix Γ^- is a generalized inverse of

$$\Gamma := \left[\begin{array}{c|ccc} \Gamma_d & & & \\ \hline & \Gamma_0 & & \\ & & \ddots & \\ & & & \Gamma_{\mu-1} \end{array}\right] \in L(\mathbb{R}^{(\mu+1)m}, \mathbb{R}^m)$$

having the properties $\Gamma\Gamma^- = I_m$ and

$$\Gamma^-\Gamma = \left[\begin{array}{c|ccc} \Gamma_d^-\Gamma_d & & & \\ \hline & \Gamma_0^-\Gamma_0 & & \\ & & \ddots & \\ & & & \Gamma_{\mu-1}^-\Gamma_{\mu-1} \end{array}\right] = \left[\begin{array}{c|cccc} \Pi_{\mu-1} & & & & \\ \hline & Q_0 & & & \\ & & \Pi_0 Q_1 & & \\ & & & \ddots & \\ & & & & \Pi_{\mu-2}Q_{\mu-1} \end{array}\right].$$

The product $K := \Gamma \begin{bmatrix} I \\ \vdots \\ I \end{bmatrix} = \begin{bmatrix} \boxed{\Gamma_d} \\ \boxed{\Gamma_0} \\ \vdots \\ \boxed{\Gamma_{\mu-1}} \end{bmatrix}$ is nonsingular. Our decomposition now means that

$$\begin{aligned} x &= \Pi_{\mu-1}x + Q_0 x + \Pi_0 Q_1 x + \cdots + \Pi_{\mu-2}Q_{\mu-1}x \\ &= [I \cdots I]\Gamma^-\Gamma \begin{bmatrix} I \\ \vdots \\ I \end{bmatrix} x = [I \cdots I] \begin{bmatrix} u \\ v_0 \\ \vdots \\ v_{\mu-1} \end{bmatrix} \end{aligned}$$

and the transformation (1.63) reads

$$\begin{bmatrix} \tilde{u} \\ \tilde{v}_0 \\ \vdots \\ \tilde{v}_{\mu-1} \end{bmatrix} = \Gamma \begin{bmatrix} u \\ v_0 \\ \vdots \\ v_{\mu-1} \end{bmatrix} = \Gamma\Gamma^-\Gamma \begin{bmatrix} I \\ \vdots \\ \vdots \\ I \end{bmatrix} x = \Gamma \begin{bmatrix} I \\ \vdots \\ \vdots \\ I \end{bmatrix} x = Kx = \tilde{x}.$$

Thus, turning from the original DAE (1.30) to the DAE in the form (1.65) means a coordinate transformation $\tilde{x} = Kx$, with a nonsingular matrix K, combined with a scaling by

$$L := [I \cdots I]\, \Gamma^-\Gamma \begin{bmatrix} \Pi_{\mu-1} & & & & \\ & Q_0 P_1 \cdots P_{\mu-1} & & & \\ & & \ddots & & \\ & & & Q_{\mu-2}P_{\mu-1} & \\ & & & & Q_{\mu-1} \end{bmatrix} \begin{bmatrix} I \\ \vdots \\ \vdots \\ I \end{bmatrix} G_\mu^{-1}.$$

L is a nonsingular matrix. Namely, $LG_\mu z = 0$ means that

$$\begin{aligned} \Pi_{\mu-1}z + Q_0 P_1 \cdots P_{\mu-1}z + \Pi_0 Q_1 P_2 \cdots P_{\mu-1}z + \cdots \\ + \Pi_{\mu-3}Q_{\mu-2}P_{\mu-1}z + \Pi_{\mu-2}Q_{\mu-1}z = 0, \end{aligned}$$