On the History of
Differential-Algebraic Equations

A Retrospective with
Personal Side Trips

Bernd Simeon

Felix-Klein-Zentrum, TU Kaiserslautern

D-67663 Kaiserslautern, Germany

`simeon@mathematik.uni-kl.de`

June 15th, 2015

1   Introduction

To write about the history of a subject is a challenge that grows with the number of pages as the original goal of completeness is turning more and more into an impossibility. With this in mind, the present article takes a very narrow approach and uses personal side trips and memories on conferences, workshops, and summer schools as the stage for some of the most important protagonists in the field and their contributions.

Completeness is thus out of the question, and instead it is my intention to provide a storyline that intersperses facts and results with background information. The latter is particularly important in teaching. In my experience students love personal stories about those who first found the theorem they are confronted with. For this reason, I hope that this work will not only be of interest for colleagues and researchers in general, but also for the next generations of motivated PhD students who choose the rich topic of differential-algebraic equations as their subject.

The paper is organized as follows. Under the headline 'The Early Days' I recall my first encounter with DAEs way back in 1987 and then go further back in time, with particular focus on the application fields in mechanics and electric circuits that finally would trigger an avalanche of research in applied mathematics and in the engineering sciences. The second section is called 'The Boom Days' and covers essentially the period from 1989 until

1996 when the DAEs had become a hot topic and attracted more and more researchers. Finally, the last section carries the title 'Consolidation' and highlights the developments of the following 10 years until 2006 when an Oberwolfach Workshop celebrated '25 Years of DAEs'.

As pointed out above, this essay does not aim at completeness. Those readers who would like to know more about the topic of DAEs and the rich oeuvre that has accumulated over the years are referred to the monographs of Brenan, Campbell & Petzold [BCP96], Griepentrog & März [GM86], Hairer & Wanner [HW96], Kunkel & Mehrmann [KM06], Lamour, März & Tischendorf [LMT13], and to the survey of Rabier & Rheinboldt [RR02].

2 The Early Days

Who are the pioneers that first studied the subject of differential-algebraic equations? And what was the motivation to look into such systems? This section starts at the end of the *'early days'* when I personally happened to learn about DAEs and then goes further back in time, arriving finally at the works of Kirchhoff [Kir47] and Lagrange [Lag88] who introduced differential equations with constraints in order to model electric circuits and mechanical systems.

2.1 First Encounter

It was the summer of 1987 when I first got in touch with the topic of DAEs. At that time, I was a math student at TU München, and I took part at a seminar on *Numerical Methods for Electric Circuit Analysis* organized by Claus Führer, Albert Gilg, and Peter Lory, under the guidance of Roland Bulirsch. Several of the student presentations in the seminar dealt with the transient analysis of electric circuits and the quest for the development of appropriate time integration methods. Since my own presentation, however, was concerned with sparsity considerations and the efficient solution of linear systems of equations, the DAEs did not really attract my attention.

One year had to pass until this attitude would eventually change. Meanwhile, Claus Führer had completed his Ph.D. at TU München and was back at the German Aerospace Center (DLR) in Oberpfaffenhofen, and he offered me an interesting topic for the Diploma Thesis, with Peter Rentrop as supervisor at TU München. The topic was concerned with the computation of

```
SUBROUTINE DDASSL (RES, NEQ, T, Y, YPRIME, TOUT, INFO, RTOL, ATOL,
+    IDID, RWORK, LRW, IWORK, LIW, RPAR, IPAR, JAC)
```

Figure 1: Calling sequence of the DASSL code [Pet82, BCP96] that has had an enormous impact on the subject of DAEs and that is still in wide use today.

quasi-stationary solutions of DAEs arising in mechanical multibody systems, with special focus on wheel-rail dynamics. In order to be able to draw on the expertise of the engineers and practitioners at DLR, I got a contract to work there as a student assistant and wrote most of the thesis at the lab in Oberpfaffenhofen.

In June 1988, a couple of weeks after I had started at DLR, Claus Führer asked me to help him with the preparation of a 3-day workshop on *Numerical Time Integration Methods for ODEs and DAEs* that was hosted by the Carl-Cranz-Gesellschaft e.V., a society that provides continuing educations and training for the engineering sciences. The main speaker of the workshop was Linda Petzold, and thus I had the great opportunity to attend her lessons and also to run simulations with the DASSL code [Pet82], see also Fig. 1.

In her talks, Linda Petzold typically began with *fully implicit systems*

$$\boldsymbol{F}(\dot{\boldsymbol{x}}, \boldsymbol{x}, t) = \boldsymbol{0} \tag{1}$$

with state variables $\boldsymbol{x}(t) \in \mathbb{R}^{n_x}$ and a nonlinear, vector-valued function $\boldsymbol{F}$ of corresponding dimension. Clearly, if the $n_x \times n_x$ Jacobian $\partial \boldsymbol{F} / \partial \dot{\boldsymbol{x}}$ is invertible, then by the implicit function theorem, it is theoretically possible to transform (1), at least locally, to an explicit system of ordinary differential equations. If $\partial \boldsymbol{F} / \partial \dot{\boldsymbol{x}}$ is singular, however, (1) constitutes the most general form of a *differential-algebraic equation.*

At that time, the DAEs were becoming a hot topic, in particular in numerical analysis, and Linda Petzold was one of the leading pioneers who set the pace and laid the foundation for what was to come in the years thereafter. In particular, the development of the DASSL code that she had started in the early 1980s [Pet82, BCP96] set a corner stone that still persists today.

Conceptually, it is intriguingly simple to replace the differential operator

d/dt in (1) by the Backward Differentiation Formula (BDF)

$$\varrho\boldsymbol{x}_{n+k} := \sum_{i=0}^{k} \alpha_i \boldsymbol{x}_{n+i} = \tau\dot{\boldsymbol{x}}(t_{n+k}) + \mathcal{O}(\tau^{k+1}) \tag{2}$$

where $\boldsymbol{x}_{n+i}$ stands for the discrete approximation of $\boldsymbol{x}(t_{n+i})$ with stepsize τ and where the α_i, $i = 0,\ldots,k$, denote the method coefficients that constitute the difference operator ϱ. Using the finite difference approximation $\varrho\boldsymbol{x}_{n+k}/\tau$ of the time derivative, the numerical solution of the DAE (1) then boils down to solving the nonlinear system

$$\boldsymbol{F}\left(\frac{\varrho\boldsymbol{x}_{n+k}}{\tau}, \boldsymbol{x}_{n+k}, t_{n+k}\right) = \boldsymbol{0} \tag{3}$$

for $\boldsymbol{x}_{n+k}$ in each time step, and this is exactly the underlying idea of DASSL.

I still recall the atmosphere of departure at that workshop in Oberpfaffenhofen, and over the following years, at various other meetings, I had the chance to become part of a scientific community in this field that was growing steadily. Below, I will come back to this point by interspersing further personal side trips.

2.2   Who Coined the Term DAEs?

Linda Petzolds's academic teacher is Bill Gear, who is widely recognized as the first mathematician of modern time who turned his attention to the field of DAEs. The first occurence of the term *Differential-Algebraic Equation* can be found in the title of Gear's paper *Simultaneous numerical solution of differential-algebraic equations* [Gea71b] from 1971, and in the same year his famous book *Numerical Initial Value Problems in Ordinary Differential Equations* [Gea71a] appeared where he already considers examples from electric circuit analysis in the form

$$\boldsymbol{E}\,\dot{\boldsymbol{x}} = \boldsymbol{\phi}(\boldsymbol{x}, t) \tag{4}$$

with singular capacitance matrix $\boldsymbol{E} \in \mathbb{R}^{n_x \times n_x}$ and right hand side function $\boldsymbol{\phi}$.

Moreover, it is also Gear who made the BDF methods popular for solving stiff ODE systems and who wrote one of the first sophisticated codes with variable order and variable stepsize, the DIFSUB routine, for this purpose.
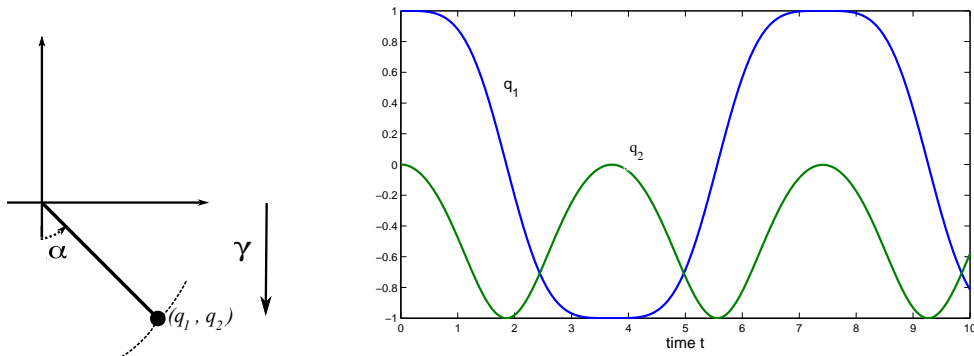
4

Figure 2: The "inevitable" pendulum.

The extension of this BDF method to linear-implicit systems (4) by means of the difference operator ϱ from (2) is straightforward and provided the first available DAE solver.

Two application fields, namely electric circuit analysis and constrained mechanical systems, are among the major driving forces for the development of DAEs. Below, this statement will be made more explicit by looking at the corresponding modeling concepts. Bill Gear had the farsightedness to very early perceive the importance of these modeling approaches for today's simulation software. During an Oberwolfach workshop in 1981, he suggested to study the *mathematical pendulum in Cartesian coordinates*

$$\ddot{q}_1 = -2q_1\lambda, \tag{5a}$$
$$\ddot{q}_2 = -\gamma - 2q_2\lambda, \tag{5b}$$
$$0 = q_1^2 + q_2^2 - 1 \tag{5c}$$

that describes the motion of a mass point with coordinates (q_1, q_2) in the plane subject to a constraint. The constraint models the massless rod of length 1 that connects the mass point to a pivot placed in the origin of the coordinate system, Fig. 2. The motion of the mass point is then determined by the gravity (parameter γ) and by the constraint forces that are expressed in terms of the unknown Lagrange multiplier λ.

The DAE (5) is an example for the Lagrange equations of the first kind that we will discuss below. By introducing velocity variables, it can be easily converted to a system of first order that fits into the class of linear-implicit DAEs (4).

5

In retrospective, the applied math community in 1981 was not ready to understand the importance of this new paradigm for modeling technical systems, and the engineering disciplines still preferred to manually transform the models to ordinary differential equations.

The notion of an index of the DAE (1) goes also back to Gear [Gea88, Gea90]. He introduced what we call today the *differentiation index*. This non-negative integer k is defined by

$k = 0$: If $\partial \boldsymbol{F}/\partial \dot{\boldsymbol{x}}$ is non-singular, the index is 0.

$k > 0$: Otherwise, consider the system of equations

$$\boldsymbol{F}(\dot{\boldsymbol{x}}, \boldsymbol{x}, t) = \boldsymbol{0},$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{F}(\dot{\boldsymbol{x}}, \boldsymbol{x}, t) = \frac{\partial}{\partial \dot{\boldsymbol{x}}}\boldsymbol{F}(\dot{\boldsymbol{x}}, \boldsymbol{x}, t)\,\boldsymbol{x}^{(2)} + \ldots = \boldsymbol{0}, \qquad (6)$$
$$\vdots$$
$$\frac{\mathrm{d}^s}{\mathrm{d}t^s}\boldsymbol{F}(\dot{\boldsymbol{x}}, \boldsymbol{x}, t) = \frac{\partial}{\partial \dot{\boldsymbol{x}}}\boldsymbol{F}(\dot{\boldsymbol{x}}, \boldsymbol{x}, t)\,\boldsymbol{x}^{(s+1)} + \ldots = \boldsymbol{0}$$

as a system in the separate dependent variables $\dot{\boldsymbol{x}}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(s+1)}$, with $\boldsymbol{x}$ and t as independent variables. Then the index k is the smallest s for which it is possible, using algebraic manipulations only, to extract an ordinary differential equation $\dot{\boldsymbol{x}} = \boldsymbol{\psi}(\boldsymbol{x}, t)$ (the underlying ODE) from (6).

Meanwhile other notions of an index have emerged, but despite its ambiguity with respect to the algebraic manipulations, the differentiation index is still the most popular and widespread tool to classify DAEs.

In the next chapter, other index concepts and their relation to the differential index will be addressed, and also more protagonists will enter the stage. This first chapter on the early days of DAEs closes now with a look at the application fields that set the ball rolling.

2.3   Kirchhoff, Weierstrass, and Kronecker

In 1847, Kirchhoff first published his *circuit laws* that describe the conservation properties of electric circuits [Kir47]. These laws consist of the current law and the voltage law, which both follow from Maxwell's equations of electro-dynamics. When these laws are applied to circuits with time-dependent behavior, the corresponding equations are typically given as a
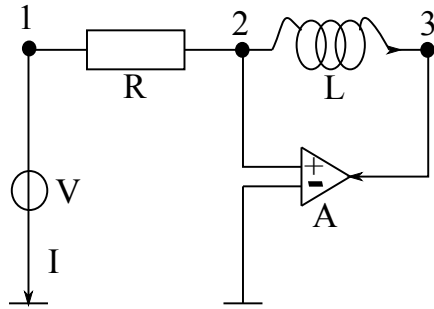
Figure 3: Differentiator circuit.

linear-implicit system (4). Often, the structure even turns out to be a linear constant coefficient DAE

$$\boldsymbol{E}\dot{\boldsymbol{x}} + \boldsymbol{H}\boldsymbol{x} = \boldsymbol{c} \tag{7}$$

with matrices $\boldsymbol{E}, \boldsymbol{H} \in \mathbb{R}^{n_x \times n_x}$ and some time-dependent source term $\boldsymbol{c}(t) \in \mathbb{R}^{n_x}$.

An example for such an electric circuit ist the *differentiator* [GHR00] shown in Fig. 3. It consists of a resistance R, an inductance L, an ideal operational amplifier $A = \infty$, and a given voltage source $V(t)$. The $n_x = 6$ unknowns read here $\boldsymbol{x} = (V_1, V_2, V_3, I, I_L, I_V)$ with voltages V_i and currents I, I_L, I_V. From Kirchhoff's laws and the properties of the amplifier and the inductance one obtains the relations

$$
\begin{aligned}
I + (V_1 - V_2)/R &= 0, \\
-(V_1 - V_2)/R + I_L &= 0, \\
-I_L + I_V &= 0, \\
V_1 &= V(t), \\
V_2 &= 0, \\
V_2 - V_3 &= L \cdot \dot{I}_L.
\end{aligned}
$$

This linear system has the form (7) with singular inductance matrix

$$\boldsymbol{E} = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & L & 0 \end{pmatrix}. \tag{8}$$

7

While the matrix $\boldsymbol{E}$ is regular in the ODE case and can be brought to the right hand side by formal inversion, it is singular in the DAE case. Weierstrass and Kronecker were in Berlin at the same time as Kirchhoff, and it is quite obvious to suppose that they knew his work [1]. Weierstrass and later Kronecker were thus inspired to study such singular systems and provided an elegant theory that is still fundamental today in order to understand the specific properties of DAEs.

We assume that the *matrix pencil* $(\boldsymbol{E}, \boldsymbol{H})$ is regular, i.e., that there exists $\mu \in \mathbb{C}$ such that the matrix $\mu\boldsymbol{E} + \boldsymbol{H}$ is regular. Otherwise, the pencil is singular, and (7) has either no or infinitely many solutions. This latter case has been first studied by Kronecker [Kro90], see also [Gan59, Cam82].

If $(\boldsymbol{E}, \boldsymbol{H})$ is a regular pencil, there exist nonsingular matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ such that

$$\boldsymbol{U}\boldsymbol{E}\boldsymbol{V} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{N} \end{pmatrix}, \quad \boldsymbol{U}\boldsymbol{H}\boldsymbol{V} = \begin{pmatrix} \boldsymbol{C} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix} \tag{9}$$

where $\boldsymbol{N}$ is a nilpotent matrix, $\boldsymbol{I}$ an identity matrix, and $\boldsymbol{C}$ a matrix that can be assumed to be in Jordan canonical form. The transformation (9) is called the *Weierstrass canonical form* [Wei68]. It is a generalization of the Jordan canonical form and contains the essential structure of the linear system (7).

In the Weierstrass canonical form (9), the singularity of the DAE is represented by the nilpotent matrix $\boldsymbol{N}$. Its degree of nilpotency, i.e., the smallest positive integer k such that $\boldsymbol{N}^k = \boldsymbol{0}$, plays a key role when studying closed-form solutions of the linear system (7) and is identical to the differentiation index of (7).

To construct a solution of (7), we introduce new variables and right hand side vectors

$$\boldsymbol{V}^{-1}\boldsymbol{x} =: \begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{z} \end{pmatrix}, \quad \boldsymbol{U}\boldsymbol{c} =: \begin{pmatrix} \boldsymbol{\delta} \\ \boldsymbol{\theta} \end{pmatrix}. \tag{10}$$

Premultiplying (7) by $\boldsymbol{U}$ then leads to the *decoupled system*

$$\dot{\boldsymbol{y}} + \boldsymbol{C}\boldsymbol{y} = \boldsymbol{\delta}, \tag{11a}$$
$$\boldsymbol{N}\dot{\boldsymbol{z}} + \boldsymbol{z} = \boldsymbol{\theta}. \tag{11b}$$

While the solution of the ODE (11a) follows by integrating and results in an expression based on the matrix exponential $\exp(-\boldsymbol{C}(t - t_0))$, the equation

---

[1] The relation of the work of Weierstrass and Kronecker to Kirchhoff's circuit laws was pointed out to me by Volker Mehrmann when we met in September 2014 during a Summer School on DAEs in Elgersburg, Germany.

(11b) for $\boldsymbol{z}$ can be solved recursively by differentiating. More precisely, it holds

$$\boldsymbol{N}\ddot{\boldsymbol{z}} + \dot{\boldsymbol{z}} = \dot{\boldsymbol{\theta}} \quad \Rightarrow \quad \boldsymbol{N}^2\ddot{\boldsymbol{z}} = -\boldsymbol{N}\dot{\boldsymbol{z}} + \boldsymbol{N}\dot{\boldsymbol{\theta}} = \boldsymbol{z} - \boldsymbol{\theta} + \boldsymbol{N}\dot{\boldsymbol{\theta}}\,.$$

Repeating the differentiation and multiplication by $\boldsymbol{N}$, we can eventually exploit the nilpotency and get

$$\boldsymbol{0} = \boldsymbol{N}^k\boldsymbol{z}^{(k)} = (-1)^k\boldsymbol{z} + \sum_{\ell=0}^{k-1}(-1)^{k-1-\ell}\boldsymbol{N}^\ell\boldsymbol{\theta}^{(\ell)}.$$

This implies the explicit representation

$$\boldsymbol{z} = \sum_{\ell=0}^{k-1}(-1)^\ell\boldsymbol{N}^\ell\boldsymbol{\theta}^{(\ell)}. \tag{12}$$

The above solution procedure illustrates several crucial points about DAEs and how they differ from ODEs. Remarkably, the linear constant coefficient case displays already these points, and thus the work of Weierstrass and Kronecker represents still the fundament of DAE theory today.

We highlight two crucial points:

(i) The solution of (7) rests on $k-1$ differentiation steps. This requires that the derivatives of certain components of $\boldsymbol{\theta}$ exist up to $\ell = k-1$. Furthermore, some components of $\boldsymbol{z}$ may only be continuous but not differentiable depending on the smoothness of $\boldsymbol{\theta}$.

(ii) The components of $\boldsymbol{z}$ are directly given in terms of the right hand side data $\boldsymbol{\theta}$ and its derivatives. Accordingly, the initial value $\boldsymbol{z}(t_0) = \boldsymbol{z}_0$ is fully determined by (12) and, in contrast to $\boldsymbol{y}_0$, cannot be chosen arbitrarily. Initial values $(\boldsymbol{y}_0, \boldsymbol{z}_0)$ where $\boldsymbol{z}_0$ satisfies (12) are called *consistent.* The same terminology applies to the initial value $\boldsymbol{x}_0$, which is consistent if, after the transformation (10), $\boldsymbol{z}_0$ satisfies (12).

Today, more than 150 years after the discoveries of Kirchhoff, electric circuit analysis remains one of the driving forces in the development of DAEs. The interplay of modeling and mathematical analysis is particularly important in this field, and the interested reader is referred to Günther & Feldmann [GF99] and März & Tischendorf [MT97] as basic works. The first simulation code that generated a model in differential-algebraic form was the SPICE package [NP73].
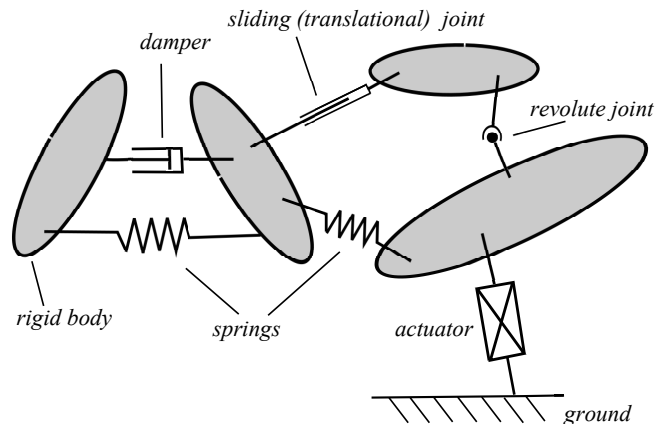
Figure 4: Sketch of a multibody system with rigid bodies and typical interconnections.

2.4 Euler and Lagrange

Even older than the DAEs arising from Kirchhoff's laws are the Euler-Lagrange equations. They were first published in Lagrange's famous work *Mécanique analytique* [Lag88] from 1788.

Consider a mechanical system that consists of rigid bodies interacting via springs, dampers, joints, and actuators, Fig. 4. The bodies possess a certain geometry and mass while the interconnection elements are massless. Let $\boldsymbol{q}(t) \in \mathbb{R}^{n_q}$ denote a vector that comprises the coordinates for position and orientation of all bodies in the system. Revolute, translational, universal, and spherical joints are examples for bondings in such a multibody system. They may constrain the motion $\boldsymbol{q}$ and hence determine its kinematics.

If constraints are present, we express the resulting conditions on $\boldsymbol{q}$ in terms of n_λ constraint equations

$$\boldsymbol{0} = \boldsymbol{g}(\boldsymbol{q}) \,. \tag{13}$$

Obviously, a meaningful model requires $n_\lambda < n_q$. The equations (13) that restrict the motion $\boldsymbol{q}$ are called *holonomic constraints*, and the rectangular matrix

$$\boldsymbol{G}(\boldsymbol{q}) := \frac{\partial \boldsymbol{g}(\boldsymbol{q})}{\partial \boldsymbol{q}} \ \in \mathbb{R}^{n_\lambda \times n_q}$$

is called the *constraint Jacobian.*

Using both the redundant position variables $\boldsymbol{q}$ and additional Lagrange multipliers $\boldsymbol{\lambda}$ to describe the dynamics leads to the *equations of constrained mechanical motion*, also called the *Lagrange equations of the first kind* or the *Euler-Lagrange equations*

$$
\begin{aligned}
\boldsymbol{M}(\boldsymbol{q})\,\ddot{\boldsymbol{q}} &= \boldsymbol{f}(\boldsymbol{q},\dot{\boldsymbol{q}},t) - \boldsymbol{G}(\boldsymbol{q})^T\boldsymbol{\lambda}\,, & \text{(14a)} \\
\boldsymbol{0} &= \boldsymbol{g}(\boldsymbol{q})\,, & \text{(14b)}
\end{aligned}
$$

where $\boldsymbol{M}(\boldsymbol{q}) \in \mathbb{R}^{n_q \times n_q}$ stands for the *mass matrix* and $\boldsymbol{f}(\boldsymbol{q},\dot{\boldsymbol{q}},t) \in \mathbb{R}^{n_q}$ for the vector of *applied and internal forces.*

The standard example for such a constrained mechanical system are the equations (5) of the mathematical pendulum. For a long time, it was common sense that the Euler-Lagrange equations should be transformed to the *state space form*, also called the *Lagrange equations of the second kind*. In case of the pendulum, this means that the Cartesian coordinates can be expressed as $q_1 = \sin\alpha$, $q_2 = -\cos\alpha$ with the angle α as minimal coordinate, Fig. 2. By inserting these relations into (5), the constraints and the Lagrange multiplier cancel, and one arrives at the second order ODE

$$
\ddot{\alpha} = -\gamma\sin\alpha \tag{15}
$$

as state space form.

It seems obvious that a state space form such as (15) constitutes a more appropriate and easier model than the differential-algebraic system (14), or (5), respectively, in redundant coordinates. In practice, however, the state space form suffers from serious drawbacks:

The analytical complexity of the constraint equations (13) makes it in various applications impossible to obtain a set of minimal coordinates that is valid for all configurations of the multibody system. Moreover, although we know from the theorem on implicit functions that such a set exists in a neighborhood of the current configuration, it might loose its validity when the configuration changes. This holds in particular for multibody systems with so-called *closed kinematic loops.*

Even more, the modeling of subsystems like electrical and hydraulic feedback controls, which are essential for the performance of modern mechanical systems, is limited. The differential-algebraic model, on the other hand, bypasses topological analysis and offers the choice of using a set of coordinates $\boldsymbol{q}$ that possess physical significance.

This reasoning in favor of the differential-algebraic model (14) became more and more widespread in the 1980s, driven by the development of sophisticated software packages, so-called *multibody formalisms*. One of the first packages that fully exploited this new way of modelling is due to Haug [Hau89].

A look at the leading software tools in the field today shows a clear picture. Some of the codes generate a differential-algebraic model whenever a constraint is present, while others try to generate a state space form as long as it is convenient. But the majority of the commercial products relies on the differential-algebraic approach as the most general way to handle complex technical applications [GB94, Sch90].

The main difference between the DAEs arising from electric circuit analysis and the DAEs that model constrained mechanical systems is the richer structure of the latter. E.g., for *conservative multibody systems*, i.e., systems where the applied forces can be written as the gradient of a potential U, the Euler-Lagrange equations (14) result from Hamilton's principle of least action

$$\int_{t_0}^{t_1} \left(T - U - \boldsymbol{g}(\boldsymbol{q})^T \boldsymbol{\lambda} \right) \mathrm{d}t \rightarrow \text{ stationary !} \tag{16}$$

where the kinetic energy possesses a representation as quadratic form

$$T(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \frac{1}{2} \dot{\boldsymbol{q}}^T \boldsymbol{M}(\boldsymbol{q}) \dot{\boldsymbol{q}} \,.$$

In the least action principle (16), we observe the fundamental Lagrange multiplier technique for coupling constraints and dynamics [Bri08]. Extensions of the multiplier technique exist in various more general settings such as dissipative systems or even inequality constraints.

The pendulum equations (5) are the example for a constrained mechanical system. Though they simply describe the motion of a single mass point, several key properties of the Euler-Lagrange equations can already be studied: the differential equations are of second order, the constraint equations are mostly nonlinear, and one observes a clear semi-explicit structure with differential variables $\boldsymbol{q}$ and algebraic variables $\boldsymbol{\lambda}$.

The Euler-Lagrange equations are of index 3 and form the prototype for a system of higher index. Index reduction techniques are thus required and in fact, already in 1972 this issue was addressed by Baumgarte [Bau72]. He observed that in (14), the Lagrange multipliers can be eliminated by

differentiating the constraints twice. The first differentiation leads to the constraints at velocity level

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\,\boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{G}(\boldsymbol{q})\,\dot{\boldsymbol{q}}. \tag{17}$$

A second differentiation step yields the *constraints at acceleration level*

$$0 = \frac{\mathrm{d}^2}{\mathrm{d}t^2}\,\boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{G}(\boldsymbol{q})\,\ddot{\boldsymbol{q}} + \boldsymbol{\kappa}(\boldsymbol{q},\dot{\boldsymbol{q}})\,, \quad \boldsymbol{\kappa}(\boldsymbol{q},\dot{\boldsymbol{q}}) := \frac{\partial \boldsymbol{G}(\boldsymbol{q})}{\partial \boldsymbol{q}}(\dot{\boldsymbol{q}},\dot{\boldsymbol{q}})\,, \tag{18}$$

where the two-form $\boldsymbol{\kappa}$ comprises additional derivative terms. The combination of the dynamic equation

$$\boldsymbol{M}(\boldsymbol{q})\,\ddot{\boldsymbol{q}} = \boldsymbol{f}(\boldsymbol{q},\dot{\boldsymbol{q}},t) - \boldsymbol{G}(\boldsymbol{q})^T\boldsymbol{\lambda}$$

with (18) results in a linear system for $\ddot{\boldsymbol{q}}$ and $\boldsymbol{\lambda}$ with the saddle point matrix

$$\begin{pmatrix} \boldsymbol{M}(\boldsymbol{q}) & \boldsymbol{G}(\boldsymbol{q})^T \\ \boldsymbol{G}(\boldsymbol{q}) & \boldsymbol{0} \end{pmatrix} \in \mathbb{R}^{(n_q+n_\lambda)\times(n_q+n_\lambda)}. \tag{19}$$

For a well-defined multibody system, this matrix is invertible in a neighborhood of the solution, and in this way, the Lagrange multiplier can be computed as a function of $\boldsymbol{q}$ and $\dot{\boldsymbol{q}}$.

However, the well-known drift-off phenomenon requires additional stabilization measures, and Baumgarte came up with the idea to combine original and differentiated constraints as

$$0 = \boldsymbol{G}(\boldsymbol{q})\,\ddot{\boldsymbol{q}} + \boldsymbol{\kappa}(\boldsymbol{q},\dot{\boldsymbol{q}}) + 2\alpha\boldsymbol{G}(\boldsymbol{q})\dot{\boldsymbol{q}} + \beta^2\boldsymbol{g}(\boldsymbol{q}) \tag{20}$$

with scalar parameters α and β. The free parameters α and β should be chosen in such a way that

$$0 = \ddot{\boldsymbol{w}} + 2\alpha\dot{\boldsymbol{w}} + \beta^2\boldsymbol{w} \tag{21}$$

becomes an asymptotically stable equation, with $\boldsymbol{w}(t) := \boldsymbol{g}(\boldsymbol{q}(t))$.

From today's perspective, the crucial point in Baumgarte's approach is the choice of the parameters. Nevertheless, it was the very beginning of a long series of works that tried to reformulate the Euler-Lagrange equations in such a way that the index is lowered while still maintaining the information of all constraint equations. For a detailed analysis of this stabilization and related techniques we refer to Ascher et al. [ACPR95, AL97].

Another – very early – stabilization of the Euler-Lagrange equations is due to Gear, Gupta & Leimkuhler [GGL85]. This formulation represents still the state-of-the-art in multibody dynamics. It uses a formulation of the equations of motion as system of first order with velocity variables $\boldsymbol{v} = \dot{\boldsymbol{q}}$ and simultaneously enforces the constraints at velocity level (17) and the position constraints (13), where the latter are interpreted as invariants and appended by means of extra Lagrange multipliers.

In this way, one obtains an enlarged system

$$
\begin{aligned}
\dot{\boldsymbol{q}} &= \boldsymbol{v} - \boldsymbol{G}(\boldsymbol{q})^T \boldsymbol{\mu}\,, \\
\boldsymbol{M}(\boldsymbol{q})\,\dot{\boldsymbol{v}} &= \boldsymbol{f}(\boldsymbol{q}, \boldsymbol{v}, t) - \boldsymbol{G}(\boldsymbol{q})^T \boldsymbol{\lambda}\,, \\
\boldsymbol{0} &= \boldsymbol{G}(\boldsymbol{q})\,\boldsymbol{v}\,, \\
\boldsymbol{0} &= \boldsymbol{g}(\boldsymbol{q})
\end{aligned}
\tag{22}
$$

with additional multipliers $\boldsymbol{\mu}(t) \in \mathbb{R}^{n_\lambda}$. A straightforward calculation shows

$$
\boldsymbol{0} = \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{g}(\boldsymbol{q}) = \boldsymbol{G}(\boldsymbol{q})\dot{\boldsymbol{q}} = \boldsymbol{G}(\boldsymbol{q})\,\boldsymbol{v} - \boldsymbol{G}(\boldsymbol{q})\boldsymbol{G}^T(\boldsymbol{q})\boldsymbol{\mu} = -\boldsymbol{G}(\boldsymbol{q})\boldsymbol{G}^T(\boldsymbol{q})\boldsymbol{\mu}
$$

and one concludes $\boldsymbol{\mu} = \boldsymbol{0}$ since $\boldsymbol{G}(\boldsymbol{q})$ is of full rank and hence $\boldsymbol{G}(\boldsymbol{q})\boldsymbol{G}^T(\boldsymbol{q})$ invertible. With the additional multipliers $\boldsymbol{\mu}$ vanishing, (22) and the original equations of motion (14) coincide along any solution. Yet, the index of the *GGL formulation* (22) is 2 instead of 3. Some authors refer to (22) also as *stabilized index-2 system.*

In the fall of 1988 – when I was finishing my master thesis at DLR Oberpfaffenhofen, Claus Führer and Ben Leimkuhler then showed that the GGL formulation in combination with a BDF discretization is basically equivalent to solving the equations of constrained mechanical motion as overdetermined system by means of a certain generalized inverse [FL89, FL91] The result became one of the most highly cited papers of those years, which demonstrates that the DAEs and their numerical analysis had attracted wide attention by then.

The last paragraphs on stabilized formulations of the Euler-Lagrange equations demonstrate that the development of theory and numerical methods for DAEs was strongly intertwined with the mathematical models. This holds for all application fields where DAEs arise. We leave this point as a loose end and turn now to what one could call the "Golden Age" of DAEs.

14

3 The Boom Days

Between 1989 and 1996, the DAEs were booming, and many groups from different fields started to explore this new research topic. Driven by the development of powerful simulation packages in the engineering sciences, the demand for efficient and robust integration methods was growing steadily while at the same time, it had become apparent that higher index problems require stabilization measures or appropriate reformulations.

This trend was reflected by a series of workshops and conferences dedicated to DAEs, and three such occasions will serve here as the stage for showcasing a – rather personal – selection of hot topics.

3.1 The Paderborn Workshops

After having finished my diploma degree, I worked for a couple of months for the DLR until the end of 1989. Sponsored by the Volkswagen Foundation, an interdisciplinary project on *Identifizierungs-, Analyse- und Entwurfsmethoden für mechanische Mehrkörpersysteme in Deskriptorform (Identification, Analysis and Design for Mechanical Multibody Systems in Descriptor Form)* gave me the opportunity to do a Ph.D. at TU München, with Peter Rentrop as supervisor. Our partners were the DLR lab in Oberpfaffenhofen with Claus Führer and Willi Kortüm and the University of Wuppertal with Peter C. Müller, who acted as coordinator of the joint project.

A part of the project plan was the organization of two workshops that should bring together the leading experts in control theory, engineering, and mathematics and thus foster the further development of DAEs. The first workshop took place in March 1992 in the Liborianum Monastery in Paderborn, and this marked the outset of a bi-annual series of workshops that would last until 2005. A recent revival meeting was organized by Sebastian Schöps and colleagues in March 2013.

Those who have attended one or more of the Paderborn Workshops recall the vivid atmosphere that was full of stimulating discussions. Confusion and misunderstandings in the communication between mathematicians and engineers happened quite often in these early days, and the distinction between a capacitor and a capacitance or the explanation of an error message *'corrector could not converge'* could result in a controversial and simultaneously entertaining discussion. The Paderborn Workshops gave me the chance to get in touch with various leading researchers in the field, thus broadening my

knowledge .

The Geneva School

At the first Paderborn meeting in 1992, Ernst Hairer gave a talk on *half-explicit Runge-Kutta methods* for semi-explicit DAEs of index 2 [Bra92, BH93]. Jointly with Christian Lubich and Michel Roche, he had written the ground-breaking monograph on *The Numerical Solution of Differential-Algebraic Equations by Runge-Kutta Methods* [HLR89] three years before. In this rich work, several new method classes, a new paradigm for the construction of convergence proofs, a new index concept, and the new RADAU5 code are presented. From then on, the Geneva School played a very strong role in the further development of DAEs and corresponding numerical methods.

The *perturbation index* as defined in [HLR89] sheds a different light on DAEs and adopts the idea of a well-posed mathematical model. While the differential index is based on successively differentiating the original DAE (1) until the obtained system can be solved for $\dot{\boldsymbol{x}}$, the perturbation index measures the sensitivity of the solutions to perturbations in the equation:

The system $\boldsymbol{F}(\dot{\boldsymbol{x}}, \boldsymbol{x}, t) = \boldsymbol{0}$ has perturbation index $k \geq 1$ along a solution $\boldsymbol{x}(t)$ on $[t_0, t_1]$ if k is the smallest integer such that, for all functions $\hat{\boldsymbol{x}}$ having a defect

$$\boldsymbol{F}(\dot{\hat{\boldsymbol{x}}}, \hat{\boldsymbol{x}}, t) = \boldsymbol{\delta}(t) \,,$$

there exists on $[t_0, t_1]$ an estimate

$$\|\hat{\boldsymbol{x}}(t) - \boldsymbol{x}(t)\| \leq c\Big(\|\hat{\boldsymbol{x}}(t_0) - \boldsymbol{x}(t_0)\| + \max_{t_0 \leq \xi \leq t} \|\boldsymbol{\delta}(\xi)\| + \ldots + \max_{t_0 \leq \xi \leq t} \|\boldsymbol{\delta}^{(k-1)}(\xi)\| \Big)$$

whenever the expression on the right hand side is sufficiently small. Note that the constant c depends only on $\boldsymbol{F}$ and on the length of the interval, but not on the perturbation $\boldsymbol{\delta}$. The perturbation index is $k = 0$ if

$$\|\hat{\boldsymbol{x}}(t) - \boldsymbol{x}(t)\| \leq c\Big(\|\hat{\boldsymbol{x}}(t_0) - \boldsymbol{x}(t_0)\| + \max_{t_0 \leq \xi \leq t} \| \int_{t_0}^{\xi} \boldsymbol{\delta}(\tau) \, \mathrm{d}\tau \| \Big) \,,$$

which is satisfied for ordinary differential equations.

If the perturbation index exceeds $k = 1$, derivatives of the perturbation show up in the estimate and indicate a certain degree of ill-posedness. E.g., if $\boldsymbol{\delta}$ contains a small high frequency term $\epsilon \sin \omega t$ with $\epsilon \ll 1$ and $\omega \gg 1$, the resulting derivatives will induce a severe amplification in the bound for $\hat{\boldsymbol{x}}(t) - \boldsymbol{x}(t)$.

Unfortunately, the differential and the perturbation index are not equivalent in general and may even differ substantially [CG95]. The story of this discovery is connected with another personal side trip in the following chapter.

The definition of the perturbation index is solely a prelude in [HLR89]. As the title says, most of the monograph deals with Runge-Kutta methods, in particular implicit ones. These are extended to linear-implicit systems $\boldsymbol{E}\dot{\boldsymbol{x}} = \boldsymbol{\phi}(\boldsymbol{x}, t)$ by assuming for a moment that the matrix $\boldsymbol{E}$ is invertible and discretizing $\dot{\boldsymbol{x}} = \boldsymbol{E}^{-1}\boldsymbol{\phi}(\boldsymbol{x}, t)$. Multiplying the resulting scheme by $\boldsymbol{E}$, one gets the method definition

$$\boldsymbol{E}\boldsymbol{X}_i = \boldsymbol{E}\boldsymbol{x}_0 + \tau \sum_{j=1}^{s} a_{ij}\boldsymbol{\phi}(\boldsymbol{X}_j, t_0 + c_j\tau), \quad i = 1, \ldots, s; \qquad (23a)$$

$$\boldsymbol{x}_1 = \left(1 - \sum_{i,j=1}^{s} b_i\gamma_{ij}\right)\boldsymbol{x}_0 + \tau \sum_{i,j=1}^{s} b_i\gamma_{ij}\boldsymbol{X}_j. \qquad (23b)$$

Here, the method coefficients are denoted by $(a_{ij})_{i,j=1}^{s}$ and $b_1, \ldots, b_s$ while $(\gamma_{ij}) = (a_{ij})^{-1}$ is the inverse of the coefficient matrix, with s being the number of stages. Obviously, (23) makes sense also in the case where $\boldsymbol{E}$ is singular.

Using *stiffly accurate methods* for differential-algebraic equations is advantageous, which becomes evident if we consider the discretization of the semi-explicit system

$$\dot{\boldsymbol{y}} = \boldsymbol{a}(\boldsymbol{y}, \boldsymbol{z}), \qquad (24a)$$

$$0 = \boldsymbol{b}(\boldsymbol{y}, \boldsymbol{z}) \qquad (24b)$$

with differential variables $\boldsymbol{y}$ and algebraic variables $\boldsymbol{z}$. The method (23) then reads

$$\boldsymbol{Y}_i = \boldsymbol{y}_0 + \tau \sum_{j=1}^{s} a_{ij}\boldsymbol{a}(\boldsymbol{Y}_j, \boldsymbol{Z}_j), \quad i = 1, \ldots, s, \qquad (25a)$$

$$0 = \boldsymbol{b}(\boldsymbol{Y}_i, \boldsymbol{Z}_i), \qquad (25b)$$

for the internal stages and

$$\boldsymbol{y}_1 \;=\; \boldsymbol{y}_0 + \tau \sum_{j=1}^{s} b_j \boldsymbol{a}(\boldsymbol{Y}_j, \boldsymbol{Z}_j), \tag{26a}$$

$$\boldsymbol{z}_1 \;=\; \left(1 - \sum_{i,j=1}^{s} b_i \gamma_{ij}\right) \boldsymbol{z}_0 + \tau \sum_{i,j=1}^{s} b_i \gamma_{ij} \boldsymbol{Z}_j \tag{26b}$$

as update for the numerical solution after one step. For stiffly accurate methods, we have $\sum_{i,j=1}^{s} b_i \gamma_{ij} = 1$ and $\boldsymbol{y}_1 = \boldsymbol{Y}_s, \boldsymbol{z}_1 = \boldsymbol{Z}_s$. The update (26) is hence superfluous and furthermore, the constraint $\boldsymbol{0} = \boldsymbol{b}(\boldsymbol{y}_1, \boldsymbol{z}_1)$ is satisfied by construction.

It is not the purpose of this article to dive further into the world of Runge-Kutta methods, but like in numerical ODEs, the rivalry between multistep methods and Runge-Kutta methods also characterizes the situation for DAEs. While Linda Petzold's DASSL code is the most prominent multistep implementation, the RADAU5 and RADAU codes [HW96, HW99] represent the one-step counter parts and have also become widespread in various applications.

The competition for the best code was a major driving force in the numerical analysis of DAEs, and from time to time those in favor of multistep methods looked also at one-step methods, e.g., in [AP91], and vice versa. Nevertheless, I would like to quote from a statement of Linda Petzold that nicely reflects the different communities: *'The BDFs are so beautiful, why would anybody consider a different method?'*

Simultaneously to the joint work with Ernst Hairer and Michel Roche, Christian Lubich investigated a different class of discretization schemes, the *half-explicit methods* [Lub89]. These methods are tailored for semi-explicit DAEs and discretize the differential equations explicitly while the constraint equations are enforced in an implicit fashion. As example, consider the Euler-Lagrange equations (14) with velocity constraint (17). The half-explicit Euler method as generic algorithm for the method class reads

$$\begin{aligned}
\boldsymbol{q}_{n+1} &= \boldsymbol{q}_n + \tau \boldsymbol{v}_n, \\
\boldsymbol{M}(\boldsymbol{q}_n)\,\boldsymbol{v}_{n+1} &= \boldsymbol{M}(\boldsymbol{q}_n)\,\boldsymbol{v}_n + \tau \boldsymbol{f}(\boldsymbol{q}_n, \boldsymbol{v}_n, t_n) - \tau \boldsymbol{G}(\boldsymbol{q}_n)^T \boldsymbol{\lambda}_n, \\
\boldsymbol{0} &= \boldsymbol{G}(\boldsymbol{q}_{n+1})\,\boldsymbol{v}_{n+1}.
\end{aligned} \tag{27}$$

Only a linear system of the form

$$\begin{pmatrix} \boldsymbol{M}(\boldsymbol{q}_n) & \boldsymbol{G}(\boldsymbol{q}_n)^T \\ \boldsymbol{G}(\boldsymbol{q}_{n+1}) & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{v}_{n+1} \\ \tau \boldsymbol{\lambda}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{M}(\boldsymbol{q}_n)\boldsymbol{v}_n + \tau \boldsymbol{f}(\boldsymbol{q}_n, \boldsymbol{v}_n, t_n) \\ \boldsymbol{0} \end{pmatrix}$$

arises here in each step. The scheme (27) forms the basis for a class of extrapolation methods [Lub89, LEN95], and also for half-explicit Runge-Kutta methods as introduced in [HLR89] and then further enhanced by Brasey & Hairer [BH93] and Arnold & Murua [AM98].

These methods have in common that only information of the velocity constraints is required. As remedy for the drift off, which grows only linearly but might still be noticeable, the following projection, which is also due to Lubich [Lub89], can be applied: Let q_{n+1} and v_{n+1} denote the numerical solution of the system, obtained by integration from consistent values q_n and v_n. Then, the projection consists of the following steps:

$$solve \quad \begin{cases} \mathbf{0} &=& \boldsymbol{M}(\tilde{\boldsymbol{q}}_{n+1})(\tilde{\boldsymbol{q}}_{n+1} - \boldsymbol{q}_{n+1}) + \boldsymbol{G}(\tilde{\boldsymbol{q}}_{n+1})^T \boldsymbol{\mu}, \\ \mathbf{0} &=& \boldsymbol{g}(\tilde{\boldsymbol{q}}_{n+1}) \end{cases} \quad for \ \tilde{\boldsymbol{q}}_{n+1}, \boldsymbol{\mu} \ (28a)$$

$$solve \quad \begin{cases} \mathbf{0} &=& \boldsymbol{M}(\tilde{\boldsymbol{q}}_{n+1})(\tilde{\boldsymbol{v}}_{n+1} - \boldsymbol{v}_{n+1}) + \boldsymbol{G}(\tilde{\boldsymbol{q}}_{n+1})^T \boldsymbol{\eta}, \\ \mathbf{0} &=& \boldsymbol{G}(\tilde{\boldsymbol{q}}_{n+1}) \, \tilde{\boldsymbol{v}}_{n+1} \end{cases} \quad for \ \tilde{\boldsymbol{v}}_{n+1}, \boldsymbol{\eta} \ (28b)$$

A simplified Newton method can be used to solve the nonlinear system (28a) while (28b) represents a linear system for $\tilde{\boldsymbol{v}}_{n+1}$ and $\boldsymbol{\eta}$ with similar structure.

The projection can also be employed for stabilizing the equations of motion with acceleration constraint (18) where the position and velocity constraints are invariants and not preserved by the time integration, see Eich [Eic93] and von Schwerin [Sch99]. Such projection methods are particularly attractive in combination with explicit ODE integrators.

DAEs and Control Theory

Control theory had and still has a considerable impact on DAEs. This holds both for interesting applications and for theoretical work. At the Paderborn Workshops, this was reflected by the participation of groups from Engineering Control and from Mathematical Control.

Peter C. Müller, organizer of the Paderborn Workshops and with degrees in mathematics and engineering perfectly suited for bringing together the different communities, was one of the first in control theory who realized that many such problems lead to DAEs in a natural way. But the traditional approach had always been to manually transform these models into ODEs, which at the time had become more and more tedious or even impossible. Classical concepts such as controlability and observability for DAEs were addressed by Müller and his co-workers already in the early 1990s and regularly presented at the Paderborn Workshops [Mu93, Mül98].

From the very beginning, Volker Mehrmann came also quite often to Paderborn. With his background from numerical linear algebra and from mathematical control theory, he brought in a completely new perspective. As he told me lately, Volker Mehrmann got first in touch with DAEs when working for the IBM Scientific Center in Heidelberg from 1988 to 1989, jointly with Peter Kunkel. They were confronted with a differential-algebraic Riccati equation

$$
\begin{aligned}
-\boldsymbol{E}(t)^T \dot{\boldsymbol{X}}(t) \boldsymbol{E}(t) \;=\;& \boldsymbol{E}(t)^T \boldsymbol{X}(t) \boldsymbol{A}(t) + \boldsymbol{A}(t)^T \boldsymbol{X}^T \boldsymbol{E}(t) + \boldsymbol{Q}(t) \quad (29) \\
& - \boldsymbol{E}(t)^T \boldsymbol{X}(t) \boldsymbol{W}(t) \boldsymbol{X}(t) \boldsymbol{E}(t)
\end{aligned}
$$

with matrices $\boldsymbol{X}(t), \boldsymbol{A}(t), \boldsymbol{Q}(t), \boldsymbol{W}(t) \in \mathbb{R}^{n \times n}$ and singular matrix $\boldsymbol{E}(t) \in \mathbb{R}^{n \times n}$. Such equations arise for example from optimal regulator problems or from optimal filters with DAE models involved, and a straightforward strategy is to rewrite the symmetric unknown matrix $\boldsymbol{X}$ into a long vector of size $n(n+1)/2$ and then to convert (29) to a DAE.

In the particular application Kunkel and Mehrmann were considering, however, this turned out to be more challenging than expected [KM90]. Even more the numerical solution of the final DAE by DASSL produced trajectories that did not match with the results of the code LIMEX, an extrapolation method that had just before been released by Deuflhard, Hairer & Zugck [DHZ87]. This surprising behavior woke the interest of Kunkel and Mehrmann for the problem class. Later on, it turned out that both codes had computed correct solutions but the equation itself admitted multiple solutions.

The Berlin School

Two and a half years after the Fall of the Wall, the first Paderborn Workshop provided also an opportunity to get in touch with Roswitha März and her co-workers from the Humboldt Universität zu Berlin. A long time before, März had already started to work on DAEs, with the book by her and Griepentrog [GM86] as the very first monograph on the topic, and over the years her projector-based analysis became the distinguishing mark of what I call here the Berlin School.

This approach is characterized by the strive for a rigorous mathematical treatment of DAEs. Following [LMT13], the projector construction can be easily illustrated by means of the constant coefficient DAE (7) that read

$$\boldsymbol{E}\dot{\boldsymbol{x}} + \boldsymbol{H}\boldsymbol{x} = \boldsymbol{c}.$$

20

While the Weierstrass canonical form (9) leads to a transformed system in new variables and is hard to compute in practice, the projector-based analysis maintains the original state variables $\boldsymbol{x}$ and proceeds as follows.

In the first step, one sets $\boldsymbol{G}_0 := \boldsymbol{E}$, $\boldsymbol{B}_0 := \boldsymbol{H}$ and determines the subspace $\mathcal{N}_0 := \ker \boldsymbol{G}_0$. For singular $\boldsymbol{G}_0$, this kernel will be non-trivial, and a projector onto $\mathcal{N}_0$ is denoted by $\boldsymbol{Q}_0$. The complementary projector is

$$\boldsymbol{P}_0 := \boldsymbol{I} - \boldsymbol{Q}_0 \,.$$

For the projectors $\boldsymbol{P}_0$ and $\boldsymbol{Q}_0$, important properties hold such as $\boldsymbol{P}_0\boldsymbol{Q}_0 = \boldsymbol{Q}_0\boldsymbol{P}_0 = \boldsymbol{0}$ and $\boldsymbol{G}_0 = \boldsymbol{G}_0(\boldsymbol{P}_0 + \boldsymbol{Q}_0) = \boldsymbol{G}_0\boldsymbol{P}_0$. The original DAE system $\boldsymbol{G}_0\dot{\boldsymbol{x}} + \boldsymbol{B}_0\boldsymbol{x} = \boldsymbol{c}$ is then equivalent to

$$\boldsymbol{G}_0\boldsymbol{P}_0\dot{\boldsymbol{x}} + \boldsymbol{B}_0(\boldsymbol{P}_0 + \boldsymbol{Q}_0)\boldsymbol{x} \;=\; \boldsymbol{c} \tag{30a}$$

$$\Leftrightarrow \quad \boldsymbol{G}_1(\boldsymbol{P}_0\dot{\boldsymbol{x}} + \boldsymbol{Q}_0\boldsymbol{x}) + \boldsymbol{B}_1\boldsymbol{x} \;=\; \boldsymbol{c} \tag{30b}$$

where

$$\boldsymbol{G}_1 := \boldsymbol{G}_0 + \boldsymbol{B}_0\boldsymbol{Q}_0, \qquad \boldsymbol{B}_1 := \boldsymbol{B}_0\boldsymbol{P}_0 \,.$$

This step is repeated in terms of

$$\boldsymbol{G}_{i+1} := \boldsymbol{G}_i + \boldsymbol{B}_i\boldsymbol{Q}_i, \qquad \boldsymbol{B}_{i+1} := \boldsymbol{B}_i\boldsymbol{P}_i \,,$$

and it can be shown that the corresponding sequence of matrices $\boldsymbol{G}_i$ in front of the derivative $\dot{\boldsymbol{x}}$ has the property

$$\operatorname{im} \boldsymbol{G}_0 \subseteq \operatorname{im} \boldsymbol{G}_1 \subseteq \ldots \subseteq \operatorname{im} \boldsymbol{G}_i \,.$$

In other words, the regularity of the leading matrix grows, and in the end $\boldsymbol{G}_i$ will become a regular matrix for some i. This is guaranteed for regular matrix pencils $(\boldsymbol{E}, \boldsymbol{H})$ where the process stops when the nilpotency index of the Weierstrass form k equals the step number i. For singular pencils, the projector-based approach provides also a means to detect and analyze the singularity.

The real power of this procedure unfolds in particular at time-variant systems

$$\boldsymbol{E}(t)\dot{\boldsymbol{x}}(t) + \boldsymbol{H}(t)\boldsymbol{x}(t) = \boldsymbol{c}(t) \,, \tag{31}$$

see [LMT13] for a recent comprehensive exposition.

Similar to the competition for the best numerical method where mostly either the BDF or the Runge-Kutta schemes have been favored by the different research groups, the projector-based analysis has contended with several
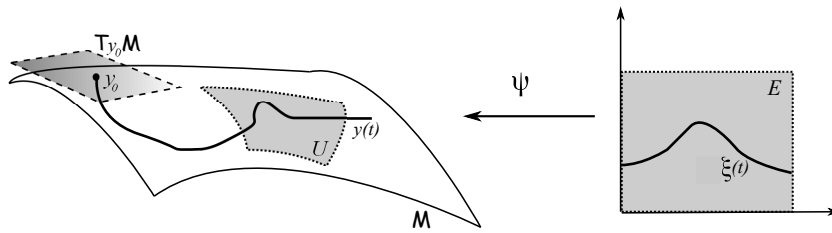
Figure 5: Manifold $\mathcal{M}$, tangent space $\mathcal{T}_y\mathcal{M}$, and local parametrization.

other approaches over the years. Among these are the derivative array technique and the interpretation of DAEs as differential equations on manifolds that will be discussed below.

3.2 The Oberwolfach Workshop in 1993

The Oberwolfach Workshop *Differential-Algebraic Equations: Theory and Applications in Technical Simulation* organized by Hans-Georg Bock, Peter Rentrop, and Werner C. Rheinboldt in June 1993 put a flashlight on the dynamic development in the field in those years. I recall the atmosphere as very stimulating and full of momentum, and for a Ph.D. student it was both encouraging – the research field was hot and one was part of a rapidly growing community – and discouraging – so many brilliant minds were already working in the field.

A particular challenge emerged during the first day as several speakers presented new time integration methods and proved their efficiency by showing results where the DASSL code was beaten when solving Andrews' squeezer, also known as the seven-body mechanism [AO86]. In this way, a benchmark was set, and during the following days a kind of horse race took place where the speakers tried to further push their integration schemes.

Differential Equations on Manifolds

Already in 1984, Werner Rheinboldt had investigated DAEs from the viewpoint of differential geometry [Rhe84]. While the approaches discussed so far are mainly inspired by differential calculus and algebraic considerations, a fundamentally different aspect comes into play by his idea of *differential equations on manifolds*.

To illustrate this approach, we consider the semi-explicit system

$$\dot{\boldsymbol{y}} = \boldsymbol{a}(\boldsymbol{y}, \boldsymbol{z}), \tag{32a}$$
$$\boldsymbol{0} = \boldsymbol{b}(\boldsymbol{y}) \tag{32b}$$

under the assumption

$$\frac{\partial \boldsymbol{b}}{\partial \boldsymbol{y}}(\boldsymbol{y}) \cdot \frac{\partial \boldsymbol{a}}{\partial \boldsymbol{z}}(\boldsymbol{y}, \boldsymbol{z}) \in \mathbb{R}^{n_z \times n_z} \quad \text{is invertible} \tag{33}$$

in a neighborhood of the solution. Clearly, (32) is of index 2 where the constraint $\boldsymbol{0} = \boldsymbol{b}(\boldsymbol{y})$, assuming sufficient differentiability, defines the manifold

$$\mathcal{M} := \{\boldsymbol{y} \in \mathbb{R}^{n_y} : \boldsymbol{b}(\boldsymbol{y}) = \boldsymbol{0}\}. \tag{34}$$

The full rank condition (33) for the matrix product $\partial \boldsymbol{b}/\partial \boldsymbol{y} \cdot \partial \boldsymbol{a}/\partial \boldsymbol{z}$ implies that the Jacobian $\boldsymbol{B}(\boldsymbol{y}) = \partial \boldsymbol{b}(\boldsymbol{y})/\partial \boldsymbol{y} \in \mathbb{R}^{n_z \times n_y}$ possesses also full rank n_z. Hence, for fixed $\boldsymbol{y} \in \mathcal{M}$, the tangent space

$$\mathcal{T}_y\mathcal{M} := \{\boldsymbol{v} \in \mathbb{R}^{n_y} : \boldsymbol{B}(\boldsymbol{y})\boldsymbol{v} = \boldsymbol{0}\} \tag{35}$$

is the kernel of $\boldsymbol{B}$ and has the same dimension $n_y - n_z$ as the manifold $\mathcal{M}$. Fig. 5 depicts $\mathcal{M}$, $\mathcal{T}_y\mathcal{M}$, and a solution of the DAE (32), which, starting from a consistent initial value, is required to proceed on the manifold.

The differential equation on the manifold $\mathcal{M}$ that is equivalent to the DAE (32) is obtained as follows: The hidden constraint

$$\boldsymbol{0} = \boldsymbol{B}(\boldsymbol{y})\boldsymbol{a}(\boldsymbol{y}, \boldsymbol{z})$$

can be solved for $\boldsymbol{z}(\boldsymbol{y})$ according to the rank condition (33) and the implicit function theorem. Moreover, for $\boldsymbol{y} \in \mathcal{M}$ it holds $\boldsymbol{a}(\boldsymbol{y}, \boldsymbol{z}(\boldsymbol{y})) \in \mathcal{T}_y\mathcal{M}$, which defines a vector field on the manifold $\mathcal{M}$ [AMR88]. Overall,

$$\dot{\boldsymbol{y}} = \boldsymbol{a}(\boldsymbol{y}, \boldsymbol{z}(\boldsymbol{y})) \quad \text{for } \boldsymbol{y} \in \mathcal{M} \tag{36}$$

represents then a differential equation on the manifold [Arn81, Rhe84].

In theory, and also computationally [Rhe96], it is possible to transform the differential equation (36) from the manifold to an ordinary differential equation in a linear space of dimension $n_y - n_z$. For this purpose, one introduces a *local parametrization*

$$\boldsymbol{\psi} : E \to \mathcal{U} \tag{37}$$

23

where E is an open subset of $\mathbb{R}^{n_y - n_z}$ and $\mathcal{U} \subset \mathcal{M}$, see Fig. 5. Such a parametrization is not unique and holds only locally in general. It is, however, possible to extend it to a family of parametrizations such that the whole manifold is covered. For $\boldsymbol{y} \in \mathcal{U}$ and local coordinates $\boldsymbol{\xi} \in E$ we thus get the relations

$$\boldsymbol{y} = \boldsymbol{\psi}(\boldsymbol{\xi}), \qquad \dot{\boldsymbol{y}} = \boldsymbol{\Psi}(\boldsymbol{\xi})\dot{\boldsymbol{\xi}}, \quad \boldsymbol{\Psi}(\boldsymbol{\xi}) := \frac{\partial \boldsymbol{\psi}}{\partial \boldsymbol{\xi}}(\boldsymbol{\xi}) \in \mathbb{R}^{n_y \times (n_y - n_z)}.$$

Premultiplying (36) by the transpose of the Jacobian $\boldsymbol{\Psi}(\boldsymbol{\xi})$ of the parametrization and substituting $\boldsymbol{y}$ by $\boldsymbol{\psi}(\boldsymbol{\xi})$, we arrive at

$$\boldsymbol{\Psi}(\boldsymbol{\xi})^T \boldsymbol{\Psi}(\boldsymbol{\xi})\dot{\boldsymbol{\xi}} = \boldsymbol{\Psi}(\boldsymbol{\xi})^T \boldsymbol{a}(\boldsymbol{\psi}(\boldsymbol{\xi}), \boldsymbol{z}(\boldsymbol{\psi}(\boldsymbol{\xi}))) \,. \tag{38}$$

Since the Jacobian $\boldsymbol{\Psi}$ has full rank for a valid parametrization, the matrix $\boldsymbol{\Psi}^T \boldsymbol{\Psi}$ is invertible, and (38) constitutes the desired ordinary differential equation in the local coordinates $\boldsymbol{\xi}$. In analogy to a mechanical system in minimal coordinates, we call (38) a *local state space form*.

The process of transforming a differential equation on a manifold to a local state space form constitutes a *push forward* operator, while the reverse mapping is called a *pull back* operator [AMR88]. It is important to realize that the previously defined concept of an index does not appear in the theory of differential equations on manifolds. Finding hidden constraints by differentiation, however, is also crucial for the classification of DAEs from a geometric point of view.

The geometrical viewpoint was also considered very early by Sebastian Reich [Rei90], but its full potential became clear only a couple of years later when the topic of geometric numerical integration emerged.

Singularly Perturbed Problems and Regularization

In the early days of DAEs, regularization was a quite popular means to convert the algebraic part into a differential equation. Motivated by physical examples such as stiff springs or parasitic effects in electric circuits, a number of authors have looked into this topic. Furthermore, it is also interesting to start with a singularly perturbed ODE, discretize it, and then to analyze the behavior of exact and numerical solution in the limit case.

To study an example for a semi-explicit system, we consider Van der Pol's equation

$$\epsilon \ddot{q} + (q^2 - 1)\dot{q} + q = 0 \tag{39}$$

with parameter $\epsilon > 0$. This is an oscillator equation with a nonlinear damping term that acts as a controller. For large amplitudes $q^2 > 1$, the damping term introduces dissipation into the system while for small values $q^2 < 1$, the sign changes and the damping term is replaced by an excitation, leading thus to a self-exciting oscillator. Introducing Liénhard's coordinates [HNW93]

$$z := q, \quad y := \epsilon\dot{z} + (z^3/3 - z),$$

we transform (39) into the first order system

$$\dot{y} = -z, \tag{40a}$$

$$\epsilon\dot{z} = y - \frac{z^3}{3} + z. \tag{40b}$$

The case $\epsilon \ll 1$ is of special interest. In the limit $\epsilon = 0$, the equation (40b) turns into a constraint and we arrive at the semi-explicit system

$$\dot{y} = -z, \tag{41a}$$

$$0 = y - \frac{z^3}{3} + z. \tag{41b}$$

In other words, Van der Pol's equation (40) in Liénhard's coordinates is an example of a *singularly perturbed system* which tends to the semi-explicit DAE (41) when $\epsilon \to 0$.

Such a close relation between a singularly perturbed system and a differential-algebraic equation is quite common and can be found in various application fields. Often, the parameter ϵ stands for an almost negligible physical quantity or the presence of strongly different time scales. Analyzing the *reduced system*, in this case (41), usually proves successful to gain a better understanding of the original perturbed equation [O'M74]. In the context of regularization methods, this relation is also exploited, but in reverse direction [Han90]. One starts with a DAE such as (41) and replaces it by a singularly perturbed ODE, in this case (40).

In numerical analysis, the derivation and study of integration schemes via a singularly perturbed ODE has been termed the *indirect approach* [HLR89] and lead to much additional insight [HLR88, LP86, Lub93], both for the differential-algebraic equation as limit case and for the stiff ODE case. A particularly interesting method class for the indirect approach are Rosenbrock methods as investigated by Rentrop, Roche & Steinebach [RRS89].
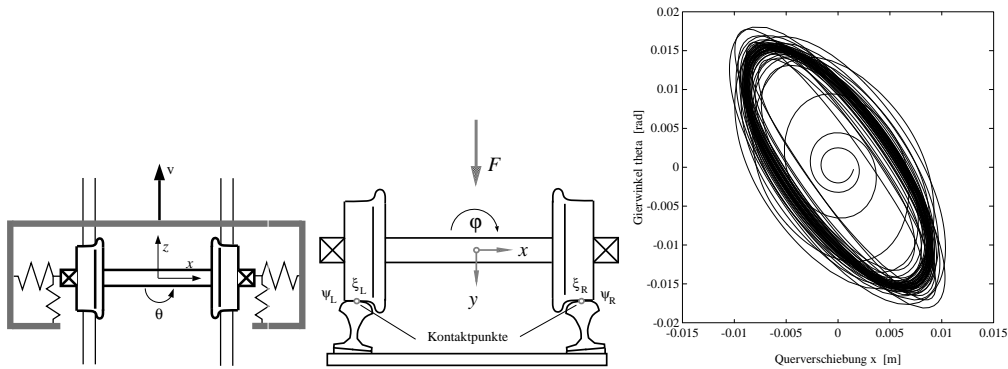
Figure 6: Wheelset on track and phase diagram of the hunting motion.

General Fully Implicit DAEs

At the Oberwolfach Workshop of 1993, I met another of the pioneers in the field of DAEs, Steve Campbell. Already in the late seventies of the last century, he had worked on singular systems of differential equations and applications in control theory, which lead to the book [Cam80]. In the first phase of the Ph.D., this book became a valuable source and inspiration for me when working on the Drazin inverse in multibody dynamics [SFR93].

Before further discussing the solution of general fully implicit DAEs (1), which was the topic of Steve Campbell's talk in Oberwolfach, it makes sense to recall the solution theory in the linear constant coefficient case While the Weierstrass transformation (9) provides the complete structural information of a linear DAE system in new coordinates and decouples the solution, the Drazin inverse represents an elegant means to express the solution in the original coordinates.

To this end, we define

$$\hat{\boldsymbol{E}} := (\mu \boldsymbol{E} - \boldsymbol{H})^{-1} \boldsymbol{E}, \quad \hat{\boldsymbol{H}} := (\mu \boldsymbol{E} - \boldsymbol{H})^{-1} \boldsymbol{H}$$

where $\mu \in \mathbb{C}$ is chosen such that the inverse of $\mu \boldsymbol{E} - \boldsymbol{H}$ exists, which is possible for a regular matrix pencil. Let $\hat{\boldsymbol{E}}$ be decomposed in Jordan canonical form, i.e.

$$\hat{\boldsymbol{E}} = \boldsymbol{T} \left[\begin{array}{cc} \boldsymbol{R} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{N} \end{array} \right] \boldsymbol{T}^{-1} \tag{42}$$

where $\boldsymbol{R}$ is associated with the non-zero eigenvalues and $\boldsymbol{N}$ is associated with the zero eigenvalues and therefore is nilpotent, as in the Weierstrass

26

canonical form (9). The Drazin inverse $\hat{\boldsymbol{E}}^D$ is defined by [Dra58]

$$\hat{\boldsymbol{E}}^D := \boldsymbol{T} \begin{pmatrix} \boldsymbol{R}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \boldsymbol{T}^{-1} \tag{43}$$

or, equivalently, by the axioms

$$\begin{array}{rlll}
\text{(D1)} & \hat{\boldsymbol{E}}\,\hat{\boldsymbol{E}}^D & = & \hat{\boldsymbol{E}}^D\,\hat{\boldsymbol{E}}\,, \\
\text{(D2)} & \hat{\boldsymbol{E}}^D\,\hat{\boldsymbol{E}}\,\hat{\boldsymbol{E}}^D & = & \hat{\boldsymbol{E}}^D\,, \\
\text{(D3)} & \hat{\boldsymbol{E}}^D\,\hat{\boldsymbol{E}}^{k+1} & = & \hat{\boldsymbol{E}}^k\,, \text{where } k \text{ is the nilpotency index of } \boldsymbol{N}\,.
\end{array}$$

The inverse $\hat{\boldsymbol{E}}^D$ always exists, is uniquely determined, and is equal to $\hat{\boldsymbol{E}}^{-1}$ for regular $\hat{\boldsymbol{E}}$. The product $\hat{\boldsymbol{E}}^D\,\hat{\boldsymbol{E}}$ is a projector which can be used to guarantee consistent initial values. Overall, for a regular matrix pencil $(\boldsymbol{E}, \boldsymbol{H})$ the homogeneous linear constant coefficient DAE $\boldsymbol{E}\dot{\boldsymbol{x}} = \boldsymbol{H}\boldsymbol{x}$ possesses the solution [Cam80]

$$\boldsymbol{x}(t) = \exp(\hat{\boldsymbol{E}}^D\hat{\boldsymbol{H}}t)\hat{\boldsymbol{E}}^D\hat{\boldsymbol{E}}\boldsymbol{x}_0\,. \tag{44}$$

The initial vector x_0 is consistent if and only if $\hat{\boldsymbol{E}}^D\hat{\boldsymbol{E}}\boldsymbol{x}_0 = \boldsymbol{x}_0$. For the inhomogeneous case see also [Cam80].

In contrast to the solution theory in the linear constant coefficient case, the treatment of fully implicit DAEs without a given internal structure is still challenging, even from today's perspective. For this purpose, Campbell [Cam93] introduced the derivative array as key concept that carries all the information of the DAE system. The derivative array is constructed from the definition of the differential index, i.e., one considers the equations

$$\begin{array}{rcl}
\boldsymbol{F}(\dot{\boldsymbol{x}}, \boldsymbol{x}, t) & = & \boldsymbol{0}, \\
\dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{F}(\dot{\boldsymbol{x}}, \boldsymbol{x}, t) = \dfrac{\partial}{\partial\dot{\boldsymbol{x}}}\boldsymbol{F}(\dot{\boldsymbol{x}}, \boldsymbol{x}, t)\,\boldsymbol{x}^{(2)} + \ldots & = & \boldsymbol{0}, \\
& \vdots & \\
\dfrac{\mathrm{d}^k}{\mathrm{d}t^k}\boldsymbol{F}(\dot{\boldsymbol{x}}, \boldsymbol{x}, t) = \dfrac{\partial}{\partial\dot{\boldsymbol{x}}}\boldsymbol{F}(\dot{\boldsymbol{x}}, \boldsymbol{x}, t)\,\boldsymbol{x}^{(k+1)} + \ldots & = & \boldsymbol{0}
\end{array} \tag{45}$$

for a DAE of index k. Upon discretization, (6) becomes an overdetermined system that can be tackled by least squares techniques. The challenge in this procedure, however, is the in general unknown index k and its determination.

Algorithms based on the derivative array are a powerful means for general unstructured DAE systems, and this holds even for the linear constant coefficient case since the computation of the Weierstrass form or the Drazin inverse are very sensitive to small perturbations and thus problematic in finite precision arithmetic. For the derivative array, in contrast, so-called staircase algorithms have been developed that rely on orthogonal matrix multiplications and are much more stable [Ben15].

The Oberwolfach Workshop of 1993 presented various other new developments that would be worthwhile for an exposition. An example is the dummy derivatives technique by Mattson and Söderlind [MS93] that provides a method to lower the index of an unstructured DAE and that is in use in today's general modelling languages.

The Flying Wheelset

Besides Andrews squeezer, there are other benchmark examples in multibody dynamics that have been established over the years. A single wheelset running on a straight track is such an example, and for a while it became well-known due to the paper [Eic93] by Edda Eich, who had discovered a strong drift-off for the formulation of index 1 and presented this result at the Oberwolfach Workshop of 1993. There is a lesson to be learnt from this example, and I like to tell the following story when teaching to Ph.D. and master students.

In 1994, one year after the *flying wheelset* had taken off, Sebastian Reich contacted me to send him the source code so that he could run some numerical tests with it. At that time, he was working on a class of stabilization methods that are related to Baumgarte's approach, jointly with Uri Ascher, H Chin and Linda Petzold [ACPR95]. The wheelset seemed a perfect example for these methods, but Sebastian Reich discovered some strange results that contradicted the theory and made him suspicious. So I received an e-mail where he described the results and questioned our Fortran code. This request made Claus Führer and me cross-check the code that we had written five years before when working on the survey paper [SFR91]. And it turned out that the constraints on acceleration level had a flaw that had been introduced when merging two blocks of output from a computer algebra program. A simple − sign was false, and after having corrected it, the drift off was drastically reduced – the wheelset had landed.

In conclusion, the cross-checking of numerical results and the exchange of

codes and benchmark problems are absolutely essential for our field in order to reproduce results and eliminate human errors.

3.3   The Oberwolfach Workshop in 1996

In some sense, the Oberwolfach Workshop in 1996, organized by Roswitha März and Linda Petzold, marks the end of the 'Boom Days'. At that time, several groups were heading into new fields such as geometric integration and partial diffferential-algebraic equations. In various respects, a solid body of knowledge had emerged by then, which is reflected by the books of Brenan, Campbell & Petzold [BCP96] and Hairer & Wanner [HW96] that both appeared in 1996. Though primarily targeting at numerical methods, both works have meanwhile become standard references on DAEs in general.

A Famous Inequality and Why One Should Never Trust Authorities

At the Oberwolfach Workshop in 1996, Steve Campbell gave a talk on the relation between the differential and the perturbation index and showed that these notions are not equivalent in general and may even differ substantially [CG95].

This surprising revelation brings me to another side trip that I love to tell master and PhD students. In 1990, Bill Gear had written the paper [Gea90] where he addressed the new perturbation index and proved the inequality

$$\mathrm{DI} \leq \mathrm{PI} \leq \mathrm{DI} + 1. \tag{46}$$

Here, DI stands for the differential index and PI for the perturbation index. I had the pleasure to attend a summer school on DAEs in Paris in 1992 where Gear talked about DAEs in general and the index notions in particular. The school had been organized by Linda Petzold, and besides her and Gear, Claus Führer and Christian Lubich were also among the speakers.

After Gear's talk, it seemed that everybody in the audience was convinced that (46) was right and another milestone in the development of DAE's had been reached. Back home in Munich, I had a master student working on the paper [Gea90] in order to prepare a seminar talk. The student was bright and repeatedly came to my office for questions about the proof of (46). In the end, we both were not able to completely follow the lines of reasoning, but I wiped away any doubts by declaring that the great Bill Gear would definitely be right. But he was not.

The counter example found by Steve Campbell is simple. It reads

$$\begin{pmatrix} 0 & y_3 & 0 \\ 0 & 0 & y_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \mathbf{0}\,. \tag{47}$$

The last equation is $y_3 = 0$, which immediately implies $y_1 = 0$ and $y_2 = 0$. Differentiating these equations once yields the underlying ordinary differential equation, and accordingly the differential index equals 1. If the right hand side, on the other hand, is perturbed by $\boldsymbol{\delta} = (\delta_1, \delta_2, \delta_3)^T$, we can compute the perturbed solution in a way similar to the derivation of (12), obtaining eventually an expression for $\hat{y}_1$ that involves the second derivative $\delta_3^{(2)}$. The perturbation index is hence 3.

The example (47) extends easily to arbitrary dimension n_y. While the perturbation index equals n_y and grows with the dimension, the differential index stays at 1. In case of semi-explicit systems, however, such an inconsistence does not arise, and both indices can be shown to be equivalent.

The bottom line of this story is clear. As we all are human, our results might be wrong and call for validation by colleagues and students. If I had really put the result (46) into question, I would have had the chance to work on a counter example on my own, at a time when I was still a PhD student. But I missed the chance since I had too much trust in authorities.

4 Consolidation

By the mid-nineties of the last century, the boom days slowly turned into a constant and stable flow of ongoing work. Furthermore, DAEs could be found all around the world in different languages and scientific contexts. To illustrate this, Fig. 7 displays the first page of a japanese text on DAEs by Naoyuki Ohsako and Masaharu Nakashima.

4.1 The NUMDIFF Conference in 1997

The series of NUMDIFF seminars goes back to the 1980s when Karl Strehmel initiated a conference format that brought eastern and western mathematicians together in the venerable city of Halle in East Germany. Concentrating on time-dependent problems and specific integration methods, NUMDIFF
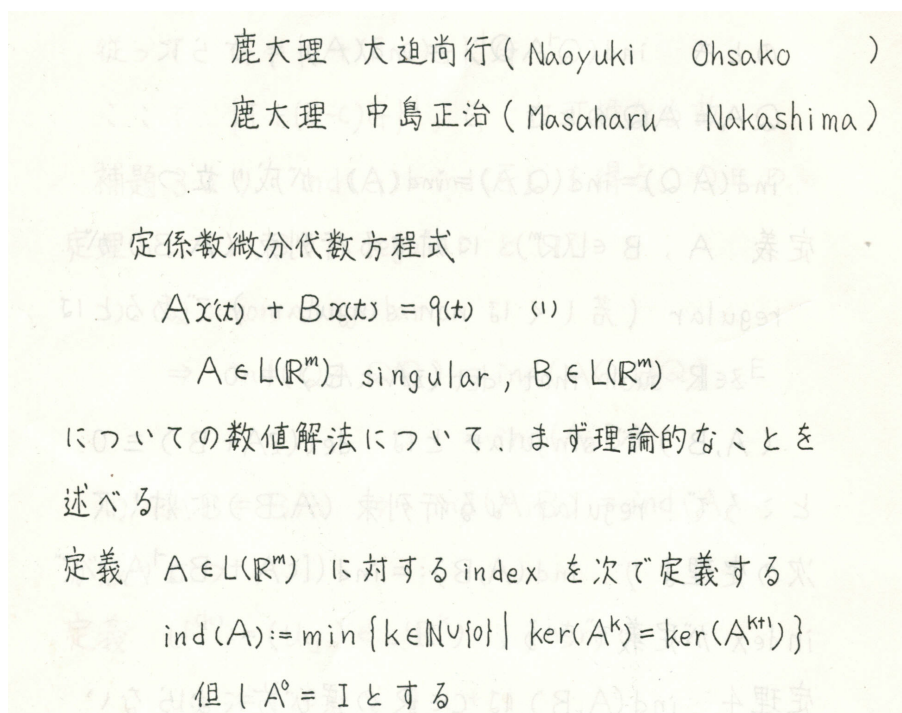
Figure 7: Title page of a japanese text on linear constant coefficient DAEs. By N. Ohsako and M. Nakashima, Dept. of Mathematics, Kagoshima University, Japan.

filled a gap and soon became a well-established conference that still takes place today.

In the early 1990s, the DAEs were also a prominent topic at the NUMD-IFF seminars, but only in 1997, when the conference moved to Alexisbad in the Harz Mountains, NUMDIFF really had a focus on DAEs and offered the stage for a new and long-lasting development. This conference marks the outset of the topic of Partial Differential Algebraic Equations (PDAEs).

The Halle group was one of the driving forces in this emerging field. An example for a linear PDAE in the unknown $\boldsymbol{u}(x,t) \in \mathbb{R}^{n_x}$ is given by

$$\boldsymbol{E}\boldsymbol{u}_t + \boldsymbol{H}\boldsymbol{u}_{xx} + \boldsymbol{C}\boldsymbol{u} = \boldsymbol{c} \tag{48}$$

where at least one of the square matrices $\boldsymbol{E}, \boldsymbol{H} \in \mathbb{R}^{n_x \times n_x}$ is singular, Lucht et al. [LSEL99]. Obviously, (48) is a generalization of the linear constant coefficient DAE (7), and one is tempted to directly transfer the already available

concepts and techniques to this problem field. However, as the theory of partial differential equations is much more heterogeneous than the one for ordinary differential equations, a general methodology for PDAEs is more than a hard task, and it is more rewarding to study special classes.

One specific aspect concerns the influence of the discretization of the spatial variable x and its derivative $\boldsymbol{u}_{xx}$, which leads to a finite-dimensional DAE in time t. This discretization has clearly an influence on the structure and may even affect the index of the resulting system. At the time when the PDAEs began to attract attention, there was much expectation that such fundamental questions could be answered in some generality. Over the years, however, it turned out that it is more advantageous to look into particular application fields. Moreover, the well-established PDE and numerical PDE communities were quite reluctant to accept the viewpoint of differential-algebraic equations and considered it as a game that is not worth the candle.

Examples for PDAEs

It requires convincing examples to demonstrate the benefits of a differential-algebraic viewpoint in the PDE context. Two such examples are sketched next.

A classical example for a PDAE is given by the Navier-Stokes equations

$$\dot{\boldsymbol{u}} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} + \tfrac{1}{\rho}\nabla p = \nu \boldsymbol{\Delta} \boldsymbol{u} + \boldsymbol{l}, \tag{49a}$$

$$0 = \nabla \cdot \boldsymbol{u} \tag{49b}$$

for the velocity field $\boldsymbol{u}(x,t)$ and the pressure $p(x,t)$ in a d-dimensional domain Ω, with mass density ρ, viscosity ν, and source term $\boldsymbol{l}(x,t)$. The second equation (49b) models the incompressibility of the fluid and defines a constraint for the velocity field. For simplification, the convection term $(\boldsymbol{u} \cdot \nabla)\boldsymbol{u}$ in (49a) can be omitted, which makes the overall problem linear and more amenable for the analysis. In an abstract notation, the resulting Stokes problem then reads

$$\dot{\boldsymbol{u}} + \mathcal{A}\boldsymbol{u} + \mathcal{B}'p = \boldsymbol{l}, \tag{50a}$$

$$\mathcal{B}\boldsymbol{u} = 0, \tag{50b}$$

with differential operators $\mathcal{A}$ and $\mathcal{B}$ expressing the Laplacian and the divergence, respectively. The notation $\mathcal{B}'$ stands for the conjugate operator of $\mathcal{B}$, which here is the gradient.

The discretization, e.g., by a Galerkin-projection

$$\boldsymbol{u}(x,t) \doteq \boldsymbol{N}(x)\boldsymbol{q}(t), \qquad p(x,t) \doteq \boldsymbol{Q}(x)\boldsymbol{\lambda}(t)$$

with ansatz functions $\boldsymbol{N}$ and $\boldsymbol{Q}$ in some finite element spaces, transforms the infinite-dimensional PDAE (50) to the DAE

$$\boldsymbol{M}\dot{\boldsymbol{q}} + \boldsymbol{A}\boldsymbol{q} + \boldsymbol{B}^T\boldsymbol{\lambda} = \boldsymbol{l}, \tag{51a}$$
$$\boldsymbol{B}\boldsymbol{q} = \boldsymbol{0}. \tag{51b}$$

While the mass matrix $\boldsymbol{M}$ and stiffness matrix $\boldsymbol{A}$ are symmetric positive definite and symmetric positive semi-definite, respectively, and easy to handle, the constraint matrix $\boldsymbol{B}$ is generated by mixing the discretizations for the velocity field and the pressure. It is well-known in mixed finite elements [BF91] that a bad choice for the discretization will either result in a rank-deficient matrix $\boldsymbol{B}$ or in a situation where the smallest singular value of $\boldsymbol{B}$ is approaching zero for a decreasing mesh size. This means that the DAE (51) may become singular or almost singular due to the spatial discretization. The famous LBB-condition by Ladyshenskaja, Babŭska, and Brezzi [BF91] gives a means to classify the discretization pairs for $\boldsymbol{u}$ and p. If the matrix $\boldsymbol{B}$ has full rank, the index of the DAE (51) is $k = 2$.

To summarize, PDEs with constraints such as the Navier-Stokes equations often feature a rich structure that should be exploited, and building on the available PDE methodology reveals interesting cross-connections with the differential-algebraic viewpoint. In this context, the abstract formulation (50) as *transient saddle point problem* defines a rather broad problem class where many application fields can be subsumed [Sim13].

By combining the state-of-the-art in DAEs with advanced PDE methodology and numerics, powerful algorithms can then be developed that break new ground. Time-space adaptivitiy for PDAEs is one such topic where many different aspects are put together in order to set up numerical schemes with sophisticated error control. The work by Lang [Lan13] defines a cornerstone in this field.

A time-space adaptive solver for the Navier-Stokes equations (49) can be constructed in the following way. Discretization in time by the implicit midpoint rule with stepsize τ yields

$$\frac{\boldsymbol{u}_{i+1} - \boldsymbol{u}_i}{\tau} + (\boldsymbol{u}_{i+\frac{1}{2}} \cdot \nabla)\boldsymbol{u}_{i+\frac{1}{2}} + \frac{1}{\rho}\nabla p_{i+1} = \nu\Delta\boldsymbol{u}_{i+\frac{1}{2}} + \boldsymbol{l}_{i+\frac{1}{2}}, \tag{52a}$$
$$\nabla \cdot \boldsymbol{u}_{i+1} = 0 \tag{52b}$$

Figure 8: Time-space adaptive solution of flow over obstacle, RE = 1000

where $\boldsymbol{u}_{i+1/2} = (\boldsymbol{u}_i + \boldsymbol{u}_{i+1})/2$, $\boldsymbol{l}_{i+1/2} = \boldsymbol{l}(x, t_i + \tau/2)$. Note that (52b) is evaluated at time t_{i+1}, which is a typical technique in DAE time integration and enforces the constraint at the next time step. The discrete pressure p_{i+1} is interpreted in the same way.

In this form, the system (52) represents a sequence of stationary nonlinear PDE problems, which is the backbone of the reverse method of lines. This method has mainly been investigated in the context of parabolic partial differential equations [Bor91, FDE$^+$06] but plays also a role in various other applications [BR13].

The key idea for adaptivity in time and space is now that, like in ODE and DAE time integration, the basic time stepping scheme (52) is combined with a second method to obtain an error estimator in time. The error estimation in space, on the other hand, is performed while solving (52) by an adaptive finite element method.

As computational example, taken from [PS08], Figure 8 shows a snapshot of the flow over an obstacle in a pipe at Reynolds number RE = 1000. Here, P_1 finite elements for both velocity field and pressure are employed, stabilized by Streamline Galerkin Least Squares [HFH89]. For the time integration, the implicit midpoint scheme (52) is combined with a simple implicit Euler step. At the bottom the current mesh is displayed and on top the vorticity of the corresponding velocity field. The adaptive algorithm captures the solution details by placing additional grid points in areas where the vorticity is high. On the other hand, unnecessary grid points in other areas are automatically removed. In this example, the adaption criterion kept the number of unknowns at around 21000 per time step.

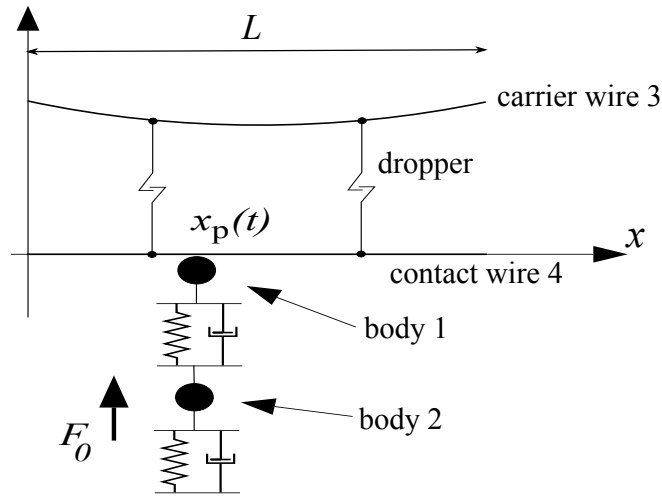As a final remark, it should be stressed that the simulation for the results

Figure 9: Pantograph and catenary, simplified benchmark problem of [AS00]

in Fig. 8 requires profound numerical skills from different fields and is not straightforward to set up. For more details on time-space adaptivity and the reverse method of lines, the reader is referred to the above references.

Pantograph and Catenary

While the Navier-Stokes equations are a PDAE system that features an explicit constraint defined over the whole domain, many PDAEs actually arise from coupling subsystems with a different level of mathematical modeling. In computational mechanics, flexible multibody systems are a typical member of this problem class. While the rigid body dynamics results in ODEs and DAEs, the inclusion of elastic or flexible bodies leads to the PDEs of elasto-dynamics where the interaction with the rigid bodies leads to additional coupling equations and constraints.

The system of pantograph and catenary [AS00, SA00] is a nice example for a flexbile multibody system and, moreover, illustrates the differential-algebraic methodology for setting up the equations of motion. The following

unknowns are used in this simplified model, Fig. 9:

$$r_1(t) : \text{vertical motion of body 1 (pantograph head)},$$
$$r_2(t) : \text{vertical motion of body 2 (pantograph base)},$$
$$w_3(x,t) : \text{vertical displacement of carrier wire},$$
$$w_4(x,t) : \text{vertical displacement of contact wire}.$$

In the first step, we neglect the constraints and consider the equations of unconstrained motion that read

$$m_1 \ddot{r}_1 = -d_1(\dot{r}_1 - \dot{r}_2) - c_1(r_1 - r_2), \tag{53a}$$
$$m_2 \ddot{r}_2 = -d_2 \dot{r}_2 + d_1(\dot{r}_1 - \dot{r}_2) - c_2 r_2 + c_1(r_1 - r_2) + F_0, \tag{53b}$$
$$\rho_3 \mathtt{A}_3 \ddot{w}_3 = -\beta_3 \dot{w}_3 + \mathtt{T}_3 w_3'' - \rho_3 \mathtt{A}_3 \gamma, \tag{53c}$$
$$\rho_4 \mathtt{A}_4 \ddot{w}_4 = -\beta_4 \dot{w}_4 + \mathtt{T}_4 w_4'' - \mathtt{E}_4 \mathtt{I}_4 w_4'''' - \rho_4 \mathtt{A}_4 \gamma. \tag{53d}$$

Here, the first two equations describe the pantograph motion with damper and spring constants d_1, d_2, c_1, c_2 and a constant force F_0 which includes the influence of the gravity. The carrier is expressed by the equation of a vibrating string with tensile force $\mathtt{T}_3$ and viscous damping factor β_3. Finally, the beam equation for the contact wire includes both a pre-stress term due to the tensile force $\mathtt{T}_4$ as well as a bending stiffness term with factor $\mathtt{E}_4 \mathtt{I}_4$. The notation of the other parameters is straightforward with $\mathtt{A}$ standing for the cross section area and γ for the gravity constant.

The next step integrates the coupling conditions where we assume bilateral contact to simplify the discussion. Contact wire and carrier are interconnected by two massless droppers with relative distances l_1 and l_2 and positions $x_{p,1}$ and $x_{p,2}$. The third constraint results from the coupling of contact wire and body 1 in the moving contact point $x_p(t)$. In strong or pointwise form, we require thus

$$w_3(x_{p,1}, t) - w_4(x_{p,1}, t) + l_1 = 0, \tag{54a}$$
$$w_3(x_{p,2}, t) - w_4(x_{p,2}, t) + l_2 = 0, \tag{54b}$$
$$w_4(x_p(t), t) - r_1(t) = 0. \tag{54c}$$

To include these constraints by appropriate Lagrange multipliers in the equations of motion (53) is a bit tricky since the constraints are formulated in isolated points of a one-dimensional continuum and result in Dirac δ-distributions in the PDEs (53c) and (53d). If one passes to a weak formulation where the δ-distribution is multiplied by a test function, however, a

well-defined model with a total of 3 discrete Lagrange multipliers is obtained that are associated with the 3 constraints (54).

Overall, the resulting model can then be written in a simular fashion as the transient saddle point problem (50) where the main difference lies in the second time derivative. More precisely, by introducing suitable operators, the pantograph and catenary model can be written as [Sim13]

$$\ddot{\boldsymbol{u}} + \mathcal{A}\boldsymbol{u} + \mathcal{B}'\boldsymbol{\lambda} = \boldsymbol{l}, \tag{55a}$$
$$\mathcal{B}\boldsymbol{u} = \boldsymbol{m}, \tag{55b}$$

where $\boldsymbol{u}$ comprises all unknown discrete and continuous displacements and $\boldsymbol{\lambda}$ stands for the Lagrange multipliers. For a related approach in elastodynamics see [Alt13].

In electrical circuit simulation, the inclusion of heating effects or semiconductors results also in PDAE models where ODEs, DAEs, and PDEs are coupled via network approaches, see, e.g., [Gün01, ABGT03].

4.2 The Oberwolfach Workshop in 2006

25 years after the workshop where Bill Gear had introduced the mathematical pendulum (5) in Cartesian coordinates, Oberwolfach offered again the showcase for the latest developments in DAEs. The organizers were Stephen Campbell, Roswitha März, Linda Petzold, and Peter Rentrop. Among the participants from all over the world was Bill Gear himself, and during the week it became evident that DAEs were now well-established in many fields.

The same year, the book by Kunkel & Mehrmann [KM06] appeared, which shed new light on topics such as boundary value problems in differential-algebraic equations and the numerical treatment of fully implicit systems (1).

Most talks at the meeting addressed the field of PDAEs, but among the other prominent topics were also optimization and optimal control problems with constraints described by DAEs , see, e.g., [CR08, KKBS04] and model order reduction for descriptor systems [RS07]. It moreover turned out that even those topics which had seemed to be mature and fully understood were still ready for surprises. An example is the properly stated leading term in linear time-variant DAEs [Mär02] where (31) is replaced by

$$\boldsymbol{E}(t)\frac{\mathrm{d}}{\mathrm{d}t}\big(\boldsymbol{D}(t)\boldsymbol{x}(t)\big) + \boldsymbol{H}(t)\boldsymbol{x}(t) = \boldsymbol{c}(t), \tag{56}$$

Figure 10: Participants of the Oberwolfach Workshop 2006, Bildarchiv des Mathematischen Forschungsinstituts Oberwolfach

together with the transversality condition

$$\ker \boldsymbol{E}(t) \oplus \operatorname{im} \boldsymbol{D}(t) = \mathbb{R}^{n_x}.$$

In this way, the matrix $\boldsymbol{D}$ precisely determines the relevant derivatives and adds additional structure to the system, which is beneficial in the analysis and also for applications in control theory.

Fig. 10 shows the participants of this memorable workshop.

At this point, the survey stops – but the story of differential-algebraic equations goes on, with many new results and entertaining stories still to be found and told.

References

[ABGT03] Ali G., Bartel A., Günther M., and Tischendorf C. (2003) Elliptic partial differential-algebraic multiphysics models in electrical network design. *Mathematical Models and Methods in Applied Sciences* 13(09): 1261–1278.

[ACPR95] Ascher U., Chin H., Petzold L., and Reich S. (1995) Stabilization of constrained mechanical systems with DAEs and invariant manifolds. *J. Mech. Struct. Machines* 23: 135–158.

[AL97] Ascher U. and Lin P. (1997) Sequential regularization methods for nonlinear higher index DAEs. *SIAM J. Scient. Comput.* 18: 160–181.

[Alt13] Altmann R. (2013) Index reduction for operator differential-algebraic equations in elastodynamics. *Z. Angew. Math. Mech. (ZAMM)* 93(9): 648–664.

[AM98] Arnold M. and Murua A. (1998) Non-stiff integrators for differential–algebraic systems of index 2. *Numerical Algorithms* 19(1-4): 25–41.

[AMR88] Abraham R., Marsden J. E., and Ratiu T. (1988) *Manifolds, Tensor Analysis, and Applications*. Springer.

[AO86]   Andrews G. C. and Ormrod M. K. (1986) Advent: A Simulation
           Program for Constrained Planar Kinematic and Dynamic Sys-
           tems. Presented at the Design Engineering Technical Conference,
           Columbus, Ohio - October 5-8, 1986. Departement of Mechanical
           Engineering, University of Waterloo, Ontario, Canada, N2L 3G1.

[AP91]   Ascher U. M. and Petzold L. R. (1991) Projected implicit Runge-
           Kutta methods for differential-algebraic equations. *SIAM J. Nu-
           mer. Anal.* 28: 1097–1120.

[Arn81]   Arnold V. I. (1981) *Ordinary Differential Equations.* MIT Press.

[AS00]   Arnold M. and Simeon B. (2000) Pantograph and catenary dy-
           namics: a benchmark problem and its numerical solution. *Appl.
           Numer. Math.* 34: 345–362.

[Bau72]   Baumgarte J. (1972) Stabilization of constraints and integrals of
           motion in dynamical systems. *Comp. Meth. in Appl. Mechanics*
           1: 1 − 16.

[BCP96]   Brenan K. E., Campbell S. L., and Petzold L. R. (1996)
           *The Numerical Solution of Initial Value Problems in Ordinary
           Differential-Algebraic Equations.* SIAM, Philadelphia.

[Ben15]   Benner P. L. P. M. V. V. M. (2015) Numerical linear algebra
           methods for linear differential-algebraic equations. *DAE Forum*
           page ???

[BF91]   Brezzi F. and Fortin M. (1991) *Mixed and Hybrid Finite Element
           Methods.* Springer, New York.

[BH93]   Brasey V. and Hairer E. (1993) Half-explicit Runge–Kutta meth-
           ods for differential-algebraic systems of index 2. *SIAM J. Nu-
           mer. Anal.* 30: 538–552.

[Bor91]   Bornemann F. A. (1991) An adaptive multilevel approach to
           parabolic equations: Ii. variable-order time discretization based
           on a multiplicative error correction. *IMPACT of Computing in
           Science and Engineering* 3(2): 93–122.

[BR13]   Bangerth W. and Rannacher R. (2013) *Adaptive finite element methods for differential equations*. Birkhäuser.

[Bra92]   Brasey V. (1992) A half-explicit method of order 5 for solving constrained mechanical systems. *Computing* 48: 191–201.

[Bri08]   Brizard A. (2008) *An Introduction to Lagrangian Mechanics*. World Scientific Publishing.

[Cam80]   Campbell S. (1980) *Singular Systems of Differential Equations*. Pitman.

[Cam82]   Campbell S. L. (1982) *Singular Systems of Differential Equations II*. Research Notes in Mathematics 61. Pitman.

[Cam93]   Campbell S. L. (1993) Least squares completions for nonlinear differential-algebraic equations. *Numer. Math.* 65: 77–94.

[CG95]   Campbell S. and Gear C. (1995) The index of general nonlinear DAEs. *Numer. Math.* 72: 173–196.

[CR08]   Callies R. and Rentrop P. (2008) Optimal control of rigid-link manipulators by indirect methods. *GAMM-Mitteilungen* 31(1): 27–58.

[DHZ87]   Deuflhard P., Hairer E., and Zugck J. (1987) One-step and extrapolation methods for differential-algebraic systems. *Numerische Mathematik* 51(5): 501–516.

[Dra58]   Drazin M. (1958) Pseudo inverses in associative rays and semigroups. *Am. Math. Mon.* 65: 506–514.

[Eic93]   Eich E. (1993) Convergence results for a coordinate projection method applied to constrained mechanical systems. *SIAM J. Numer. Anal.* 30(5): 1467–1482.

[FDE+06] Franzone P. C., Deuflhard P., Erdmann B., Lang J., and Pavarino L. F. (2006) Adaptivity in space and time for reaction-diffusion systems in electrocardiology. *SIAM Journal on Scientific Computing* 28(3): 942–962.

[FL89]   Führer C. and Leimkuhler B. (1989) Formulation and Numeri-
            cal Solution of the Equations of Constrained Mechanical Motion.
            Technical Report DFVLR-FB 89-08, Deutsche Forschungs- und
            Versuchsanstalt für Luft- und Raumfahrt (DFVLR), D-5000 Köln
            90.

[FL91]   Führer C. and Leimkuhler B. (1991) Numerical solution of
            differential-algebraic equations for constrained mechanical mo-
            tion. *Numer. Math.*. 59: 55–69.

[Gan59]   Gantmacher F. (1959) *Matrizenrechnung, Teil 2.* VEB Deutscher
            Verlag der Wissenschaften, Berlin.

[GB94]   Garcia de Jalón J. and Bayo E. (1994) *Kinematic and Dynamic
            Simulation of Multibody Systems.* Springer.

[Gea71a]   Gear C. (1971) *Numerical Initial Value Problems in Ordinary
            Differential Equations.* Prentice-Hall.

[Gea71b]   Gear C. (1971) Simultaneous numerical solution of differential-
            algebraic equations. *IEEE Trans. Circuit Theory* CT-18(1): 89 –
            95.

[Gea88]   Gear C. W. (1988) Differential-algebraic equation index transfor-
            mation. *SIAM J. Sci. & Statist. Comp.* 9: 39–47.

[Gea90]   Gear C. W. (1990) Differential-algebraic equations, indices, and
            integral algebraic equations. *SIAM J. Numer. Anal.* 27: 1527–
            1534.

[GF99]   Günther M. and Feldmann U. (1999) CAD based electric circuit
            modeling in industry I: mathematical structure and index of net-
            work equations. *Surv. Math. Ind.* 8: 97–129.

[GGL85]   Gear C., Gupta G., and Leimkuhler B. (1985) Automatic integra-
            tion of the Euler-Lagrange equations with constraints. *J. Comp.
            Appl. Math.* 12 & 13: 77–90.

[GHR00]   Günther M., Hoschek M., and Rentrop P. (2000) Differential-
            algebraic equations in electric circuit simulation. *Int. J. Electron.
            Commun.* 54: 101–107.

[GM86]   Griepentrog E. and März R. (1986) *Differential-Algebraic Equations and Their Numerical Treatment*. Teubner-Texte zur Mathematik No. 88. BSB B.G. Teubner Verlagsgesellschaft, Leipzig.

[Gün01]   Günther M. (2001) *Partielle differential-algebraische Systeme in der numerischen Zeitbereichsanalyse elektrischer Schaltungen*. VDI-Verlag, Reihe 20, Düsseldorf.

[Han90]   Hanke M. (1990) On the regularization of index 2 differential-algebraic equations. *J. Math. Anal. Appl.* 151(1): 236–253.

[Hau89]   Haug E. (1989) *Computer-Aided Kinematics and Dynamics of Mechanical Systems*. Allyn and Bacon, Boston.

[HFH89]   Hughes T. J., Franca L. P., and Hulbert G. M. (1989) A new finite element formulation for computational fluid dynamics: Viii. the galerkin/least-squares method for advective-diffusive equations. *Computer Methods in Applied Mechanics and Engineering* 73(2): 173–189.

[HLR88]   Hairer E., Lubich C., and Roche M. (1988) Error of runge-kutta methods for stiff problems studied via differential algebraic equations. *BIT Numerical Mathematics* 28(3): 678–700.

[HLR89]   Hairer E., Lubich C., and Roche M. (1989) *The Numerical Solution of Differential-Algebraic Equations by Runge-Kutta Methods*. Lecture Notes in Mathematics Vol. 1409. Springer, Heidelberg.

[HNW93]   Hairer E., Nørsett S., and Wanner G. (1993) *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer-Verlag, Berlin.

[HW96]   Hairer E. and Wanner G. (1996) *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*. Springer-Verlag, Berlin.

[HW99]   Hairer E. and Wanner G. (1999) Stiff differential equations solved by Radau methods. *J. Comp. Appl. Math.* 111: 93–111.

[Kir47]   Kirchhoff G. (1847) Ueber die auflösung der gleichungen, auf welche man bei der untersuchung der linearen vertheilung galvanischer ströme geführt wird. *Annalen der Physik* 148(12): 497–508.

[KKBS04]   Körkel* S., Kostina E., Bock H. G., and Schlöder J. P. (2004) Numerical methods for optimal control problems in design of robust optimal experiments for nonlinear dynamic processes. *Optimization Methods and Software* 19(3-4): 327–338.

[KM90]   Kunkel P. and Mehrmann V. (1990) Numerical solution of differential algebraic riccati equations. *Linear algebra and its applications* 137: 39–66.

[KM06]   Kunkel P. and Mehrmann V. (2006) *Differential-Algebraic Equations – Analysis and Numerical Solution*. EMS Publishing House.

[Kro90]   Kronecker L. (1890) Algebraische Reduktion der Schaaren bilinearer Formen. *Akademie der Wissenschaften Berlin* III: 141–155.

[Lag88]   Lagrange J. L. (1788) *Méchanique analytique*. Libraire chez la Veuve Desaint, Paris.

[Lan13]   Lang J. (2013) *Adaptive multilevel solution of nonlinear parabolic PDE systems: theory, algorithm, and applications*, volume 16. Springer Science & Business Media.

[LEN95]   Lubich C., Engstler C., and Nowak U. andPöhle U. (1995) Numerical integration of constrained mechanical systemsusing MEXX. *Mech. Struct. Mach.* 23: 473–495.

[LMT13]   Lamour R., März R., and Tischendorf C. (2013) *Differential-Algebraic Equations: A Projector Based Analysis (Differential-Algebraic Equations Forum)*. Springer.

[LP86]   Lötstedt P. and Petzold L. (1986) Numerical Solution of Nonlinear Differential Equations with Algebraic Constraints I: Convergence Results for BDF. *Math.Comp.* 46: 491–516.

[LSEL99]   Lucht W., Strehmel K., and Eichler-Liebenow C. (1999) Indexes and special discretization methods for linear partial differential algebraic equations. *BIT Numerical Mathematics* 39(3): 484–512.

[Lub89]   Lubich C. (1989) h^2 extrapolation methods for differential-algebraic equations of index-2. *Impact Comp. Sci. Eng.* 1: 260–268.

[Lub93]   Lubich C. (1993) Integration of stiff mechanical systems by Runge-Kutta methods. *ZAMP* 44: 1022–1053.

[März02]   März R. (2002) Differential algebraic systems anew. *Applied Numerical Mathematics* 42(1): 315–335.

[MS93]   Mattson S. and Söderlind G. (1993) Index reduction in differential-algebraic equations using dummy derivatives. *SIAM Sc.Stat.Comp.* (3): 677–692.

[MT97]   März R. and Tischendorf C. (1997) Recent results in solving index-2 differential-algebraic equations in circuit simulation. *SIAM J. Sci. Comp.* 18: 139–159.

[Mu93]   Muller P. C. (1993) Stability of linear mechanical systems with holonomic constraints. *Applied Mechanics Reviews* 46(11S): S160–S164.

[Mül98]   Müller P. C. (1998) Stability and optimal control of nonlinear descriptor systems: A survey. *Applied Mathematics and Computer Science* 8: 269–286.

[NP73]   Nagel L. W. and Pederson D. (Apr 1973) Spice (simulation program with integrated circuit emphasis). Technical Report UCB/ERL M382, EECS Department, University of California, Berkeley.

[O'M74]   O'Malley R. E. (1974) *Introduction to Singular Perturbations.* Academic Press, New York.

[Pet82]   Petzold L. (1982) A description of DASSL: A differential/algebraic system solver. In *Proc. 10th IMACS World Congress, August 8-13 Montreal 1982.*

[PS08]   Plinninger T. and Simeon B. (2008) *Adaptivity in Space and Time for Solving Transient Problems in COMSOL.* Proceedings COMSOL Conference Hannover 2008.

[Rei90]   Reich S. (1990) On a geometric interpretation of daes. *Circ. Syst., Sig. Processing* 9: 367–382.

[Rhe84]   Rheinboldt W. (1984) Differential - algebraic systems as differential equations on manifolds. *Math. Comp.* 43(168): 2473–482.

[Rhe96]   Rheinboldt W. (1996) Manpak: A set of algorithms for computations on implicitly defined manifolds. *Computers and Math. Applic.* 32: 15–28.

[RR02]   Rabier P. and Rheinboldt W. (2002) Theoretical and numerical analysis of differential-algebraic equations. In Ciarlet P. and Lions J. (eds) *Handbook of Numerical Analysis, Volume VIII*. Elsevier, Amsterdam.

[RRS89]   Rentrop P., Roche M., and Steinebach G. (1989) The application of Rosenbrock-Wanner type methods with stepsize control in differential-algebraic equations. *Num. Math.* 55: 545–563.

[RS07]   Reis T. and Stykel T. (2007) Stability analysis and model order reduction of coupled systems. *Mathematical and Computer Modelling of Dynamical Systems* 13(5): 413–436.

[SA00]   Simeon B. and Arnold M. (2000) Coupling DAE's and PDE's for simulating the interaction of pantograph and catenary. *Math. and Computer Model. of Systems* 6: 129–144.

[Sch90]   Schiehlen W. (ed) (1990) *Multibody System Handbook*. Springer, Heidelberg.

[Sch99]   Schwerin R. (1999) *Multibody System Simulation*. Springer, Berlin.

[SFR91]   Simeon B., Führer C., and Rentrop P. (1991) Differential-algebraic equations in vehicle system dynamics. *Surv. Math. Ind.* 1: 1–37.

[SFR93]   Simeon B., Führer C., and Rentrop P. (1993) The Drazin inverse in multibody system dynamics. *Numer. Math.* 64: 521–539.

[Sim13]   Simeon B. (2013) *Computational flexible multibody dynamics: a differential-algebraic approach.* Springer Science & Business Media.

[Wei68]   Weierstrass K. (1868) Zur theorie der bilinearen und quadratischen formen. *Monatsber. Akad. Wiss. Berlin* pages 310–338.