Mathematisches Forschungsinstitut Oberwolfach

Report No. 18/2006

Differential-Algebraic Equations

Organised by
Stephen L. Campbell (Raleigh)
Roswitha März (Berlin)
Linda R. Petzold (Santa Barbara)
Peter Rentrop (München)

April 16th – April 22nd, 2006

Abstract. Differential-Algebraic Equations (DAE) are today an independent field of research, which is gaining in importance and becoming of increasing interest for applications and mathematics itself. This workshop has drawn the balance after about 25 years investigations of DAEs and the research aims of the future were intensively discussed.

Mathematics Subject Classification (2000): 34-XX, 35-XX, 49-XX, 65-XX.

Introduction by the Organisers

The topic of Differential Algebraic Equations (DAEs) began to attract significant research interest in applied and numerical mathematics in the early 1980's. Today, a quarter of a century later, DAEs are an independent field of research, which is gaining in importance and becoming of increasing interest for both applications and mathematical theory.
This Oberwolfach workshop brought together 48 experts in applied mathematics, among them, on the one hand, some who have already influenced and formed the developments of the field, and on the other hand, some very young researchers who have shown outstanding creativity and competence in connection with their PhD theses and thus raise great hopes for further advances.
The 16 female and 32 male scientists came from 13 countries to meet and work together in the wonderful, unique Oberwolfach atmosphere, which stimulated a fruitful and pleasant collaboration.
The schedule comprised a total of 34 presentations, 18 of which were arranged into 14 survey lectures (some of them with more than one speaker) offering a broader

treatment of a particular subject. 16 shorter contributions supplemented the scientific programmme. The areas can be classified (of course with large overlap) into 4 groups:

- abstract differential algebraic systems, coupled systems, partial differential algebraic systems;
- analysis of (ordinary) differential algebraic equations and application of numerical methods to problems having new mathematical complexity;
- innovative and improved numerical integration methods to solve highly complex application problems;
- optimization with constraints described by DAEs and control problems concerning DAEs.

The broad range of these areas and the diversity of the participants stimulated fruitful discussions between the different branches and gave rise to new contacts and collaborations. A considerable gain in knowledgde and progress became obvious, which includes the formulation of open questions and challenges for the future.

We are grateful to the Mathematisches Forschungsinstitut Oberwolfach for providing an inspiring setting for this workshop.

Workshop: Differential-Algebraic Equations

Table of Contents

Abstracts

DAEs at work: Industrial multibody system simulation

Martin Arnold

For more than 15 years, the dynamical simulation of constrained mechanical systems has been one of the central applications of DAE methods in engineering. In the present contribution we summarize shortly the historical background and some current developments in this field.

Following a network approach for model setup [16], the state of complex multibody systems is described by a vector $q(t) \in \mathbb{R}^{n_q}$ of in general redundant position and orientation coordinates and the corresponding velocities $v(t) := \dot{q}(t)$. Redundant coordinates have to satisfy n_λ holonomic constraints

$$0 = g(q(t), t) \tag{1}$$

that are coupled to the dynamical equations

$$M(q,t)\ddot{q}(t) = f(q, \dot{q}, t) - G^\top(q,t)\lambda \tag{2}$$

by constraint forces $-G^\top\lambda$ with $G(q,t) := (\partial g/\partial q)(q,t)$ and Lagrangian multipliers $\lambda(t) \in \mathbb{R}^{n_\lambda}$, see [18]. Eqs. (1) and (2) together form the equations of motion, a second order DAE with the symmetric positive definite mass matrix M and the force vector f. If G has full rank, the differential index and the perturbation index of (1)/(2) are $\nu = 3$, see [11, 14]. Index reduction is necessary to guarantee the robust and efficient time integration of the equations of motion.

It is important to note that the constraints (1) result from the use of redundant coordinates $q(t)$. Locally, the constraints may always be avoided using a minimal set of generalized coordinates $\bar{q}(t) \in \mathbb{R}^{n_q - n_\lambda}$. Coordinate partitioning [20] is a numerical approach to split the vector $q(t)$ of redundant coordinates into $n_q - n_\lambda$ independent coordinates $\bar{q}(t)$ and n_λ dependent ones.

With coordinate partitioning, the dynamical simulation of constrained mechanical systems could be based on any suitable ODE solver. From the engineer's viewpoint, the use of DAE methods is attractive only if they result in a faster and more robust time integration of the equations of motion. For the acceptance of DAE methods in this field it is essential to provide a convincing physical interpretation of common DAE techniques like index reduction or projection and to have robust DAE solvers that are prepared to be used by non-experts.

The use of modern DAE methods in industrial multibody system simulation was inspired by the work of Führer [10, 11, 12] who applied index reduction by differentiation, a very formal technique from DAE theory, to derive the (hidden) *constraints on the level of velocity coordinates*

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} g(q(t), t) = \frac{\partial g}{\partial q}(q,t)\,\dot{q}(t) + \frac{\partial g}{\partial t}(q,t) = G(q,t)\,v + g_t(q,t) \tag{3}$$

that are combined with (1)/(2) in an analytically equivalent second order index-2 DAE which is known today as *stabilized index-2 formulation* (or Gear–Gupta–Leimkuhler formulation [12, 13]) of the equations of motion. With ODASSL [11], a specially adapted version of the BDF solver DASSL [8], the stabilized index-2 formulation of the equations of motion in standard form (1)/(2) may be solved very efficiently.

In industrial applications, the model equations of multibody systems are not restricted to this rather simple standard form (1)/(2). An important extension are rigid body contact conditions that may be formulated efficiently using parametrizations of the surfaces of the bodies being in contact. Geometric considerations show that the contact point coordinates $s = s(q, t)$ are implicitly defined by additional algebraic equations $0 = c(q, s, t)$, see [10], which have to be appended to (1)/(2). The stabilized index-2 formulation has been extended to these more general model equations [1, 6] by an approach that is closely related to the general DAE index reduction concept according to Kunkel and Mehrmann, see [5].

Today, the stabilized index-2 formulation combined with DASSL / ODASSL or with the implicit Runge–Kutta solver RADAU5, see [14], is one of the standard approaches to the time integration of constrained mechanical systems in industrial multibody system simulation [2]. Special techniques for Jacobian approximation and Jacobian update have been developed that exploit the structure of large scale multibody system models ($n_q \gg 100$) to reduce the computing time by 80% and more [3]. With these adapted solvers, large multibody system models like detailed full vehicle models in automotive and railway engineering may still be handled efficiently ($n_q = 100 \ldots 1000$, $n_\lambda = 10 \ldots 50$).

However, in high-end applications with thousands of degrees of freedom also these adapted solvers show a dramatical loss of efficiency. Typical examples are the dynamical simulation of combustion engines with chain drives [15] and multibody system models of vehicles or vehicle components that move along large elastic structures like, e.g., a heavy truck that crosses a bridge or the pantograph head of a high-speed train that moves along the overhead equipment [7, 19].

Often, these large scale problems show a clear modular structure that can be exploited in the dynamical simulation coupling, e.g., two or more specialized simulation tools in a co-simulation framework [17] or using small stepsizes for (hopefully) low dimensional subsystems with small time constants and much larger stepsizes for the remaining part of the model (multi-rate methods [15]). For a large class of these *modular* time integration methods, stability and convergence may be studied by techniques from DAE theory. Following the ideas of the convergence analysis for one-step methods applied to semi-explicit index-1 DAEs [9], a contractivity condition is given that is necessary for stability and convergence of modular time integration [4].

References

[1] M. Arnold. *Zur Theorie und zur numerischen Lösung von Anfangswertproblemen für differentiell-algebraische Systeme von höherem Index.* Fortschritt-Berichte VDI Reihe 20, Nr. 264. VDI–Verlag, Düsseldorf, 1998.

[2] M. Arnold. Simulation algorithms and software tools. Accepted for publication in: G. Mastinu and M. Plöchl, editors, *Road and Off-Road Vehicle System Dynamics Handbook*. Taylor & Francis, London, 2006, to appear.
[3] M. Arnold, A. Fuchs, and C. Führer. Efficient corrector iteration for DAE time integration in multibody dynamics. *Comp. Meth. Appl. Mech. Eng.*, 2006, in print.
[4] M. Arnold and M. Günther. Preconditioned dynamic iteration for coupled differential-algebraic systems. *BIT Numerical Mathematics*, 41:1–25, 2001.
[5] M. Arnold, V. Mehrmann, and A. Steinbrecher. Index reduction in industrial multibody system simulation. Technical Report IB 532–01–01, DLR German Aerospace Center, Institute of Aeroelasticity, Vehicle System Dynamics Group, 2001.
[6] M. Arnold and H. Netter. Ein modifizierter Korrektor für die stabilisierte Integration differential-algebraischer Systeme mit von Hessenbergform abweichender Struktur. Technical Report IB 515–93–03, DLR, D-5000 Köln 90, 1993.
[7] M. Arnold and B. Simeon. Pantograph and catenary dynamics: a benchmark problem and its numerical solution. *Applied Numerical Mathematics*, 34:345–362, 2000.
[8] K.E. Brenan, S.L. Campbell, and L.R. Petzold. *Numerical solution of initial–value problems in differential–algebraic equations*. SIAM, Philadelphia, 2nd edition, 1996.
[9] P. Deuflhard, E. Hairer, and J. Zugck. One–step and extrapolation methods for differential–algebraic systems. *Numer. Math.*, 51:501–516, 1987.
[10] E. Eich-Soellner and C. Führer. *Numerical Methods in Multibody Dynamics*. Teubner–Verlag, Stuttgart, 1998.
[11] C. Führer. Differential-algebraische Gleichungssysteme in mechanischen Mehrkörpersystemen. Theorie, numerische Ansätze und Anwendungen. PhD thesis, TU München, Mathematisches Institut und Institut für Informatik, 1988.
[12] C. Führer and B. Leimkuhler. Numerical solution of differential–algebraic equations for constrained mechanical motion. *Numer. Math.*, 59:55–69, 1991.
[13] C.W. Gear, B. Leimkuhler, and G.K. Gupta. Automatic integration of Euler–Lagrange equations with constraints. *J. Comp. Appl. Math.*, 12&13:77–90, 1985.
[14] E. Hairer, Ch. Lubich, and M. Roche. *The numerical solution of differential-algebraic systems by Runge–Kutta methods*. Lecture Notes in Mathematics, 1409. Springer–Verlag, Berlin Heidelberg New York, 1989.
[15] G. Hippmann, M. Arnold, and M. Schittenhelm. Efficient simulation of bush and roller chain drives. In J.M. Goicolea, J. Cuadrado, and J.C. García Orden, editors, *Proc. of Multibody Dynamics 2005 (ECCOMAS Thematic Conference)*, Madrid, Spain, 2005.
[16] M. Hoschek, P. Rentrop, and Y. Wagner. Network approach and differential-algebraic systems in technical applications. *Surveys on Math. in Industry*, 9:49–75, 1999.
[17] W. Kortüm, W.O. Schiehlen, and M. Arnold. Software tools: From multibody system analysis to vehicle system dynamics. In H. Aref and J.W. Phillips, editors, *Mechanics for a New Millennium*, pages 225–238, Dordrecht, 2001. Kluwer Academic Publishers.
[18] R.E. Roberson and R. Schwertassek. *Dynamics of Multibody Systems*. Springer–Verlag, Berlin Heidelberg New York, 1988.
[19] A. Veitl and M. Arnold. Coupled simulation of multibody systems and elastic structures. In J.A.C. Ambrósio and W.O. Schiehlen, editors, *Advances in Computational Multibody Dynamics*, pages 635–644, IDMEC/IST Lisbon, Portugal, 1999.
[20] R.A. Wehage and E.J. Haug. Generalized coordinate partitioning for dimension reduction in analysis of constrained dynamic systems. *J. Mech. Design*, 104:247–255, 1982.

PDA-Models in Chip Design — Wavelet-based Integration

Andreas Bartel
(joint work with Stephanie Knorr)

1. Introduction

The usual modelling of electric circuits yields systems of differential-algebraic equations (DAEs). Due to down-scaling, secondary effects become more and more important (e.g. [1, 2, 6]): for instance, thermal-conduction, transmission line phenomena and complex semiconductor behaviour, plus additionally inherent multi-scales of signals. Here more sophisticated models enrich the DAE-description by spatial systems, which results in a partial differential-algebraic equation (PDAE) to include adequately down-scaling effects. Roughly speaking, there are three classes of models: (1) refined networks, where network elements are replaced by a spatial description of the underlying electric effect; (2) multiphysics, where additional quantities are introduced; and (3) multirate, where time scales are decoupled by multiple time variables.

We address here the last case, and investigate the detection of steep gradients in heterogeneous signal structures (digital-like plus analog) via wavelets.

2. Multidimensional Signal Model

We are interested in computing limit cycles for circuits with widely separated time scales. Here, these problems are faced by the introduction of a corresponding variable for each occurring scale [3]. The resulting multidimensional representation of a signal yields then a *multivariate function (MVF)*. We illustrate this for a 2-tone quasiperiodic signal x, which is transferred to its MVF $\hat{x}$ as follows:

$$x(t) = \sin\left(\frac{2\pi}{T_1}\,t\right)\sin\left(\frac{2\pi}{T_2}\,t\right) \quad \rightsquigarrow \quad \hat{x}(t_1,t_2) = \sin\left(\frac{2\pi}{T_1}\,t_1\right)\sin\left(\frac{2\pi}{T_2}\,t_2\right).$$

In the multidimensional description the time scales are decoupled. In this example, the MVF is periodic in each coordinate direction and can be resolved with only few grid points over the rectangle of the periodicities $[0,T_1]\times[0,T_2]$. The more the time scales differ ($T_1 \gg T_2$), the more efficient the multidimensional approach becomes, since the structure of the MVF is independent from the ratio T_1/T_2 in contrast to the original x, which can be reconstructed via $x(t) = \hat{x}(t,t)$.

Applying the multidimensional signal model to differential-algebraic network equations leads to *multirate partial differential-algebraic equations (MPDAEs)*:

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{q}(\mathbf{x}(t)) = \mathbf{f}\big(\mathbf{b}(t),\mathbf{x}(t)\big) \quad \rightsquigarrow \quad \frac{\partial \mathbf{q}(\hat{\mathbf{x}})}{\partial t_1} + \frac{\partial \mathbf{q}(\hat{\mathbf{x}})}{\partial t_2} = \mathbf{f}\big(\hat{\mathbf{b}}(t_1,t_2),\hat{\mathbf{x}}(t_1,t_2)\big)$$

with MVFs $\hat{\mathbf{x}}$ of the unknown node potentials and branch currents and $\hat{\mathbf{b}}$ representing input signals; the charges and fluxes are described by $\mathbf{q}$.

In analogy to the theory of an underlying ODE, a structural analysis of the MPDAE [5] revealed the characterisation as a PDE-system restricted to a manifold.

The hyperbolic type of the inherent PDE allows us to formulate a characteristic system. We are left with systems of differential-algebraic equations for $\bar{\mathbf{x}}_c$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{q}(\bar{\mathbf{x}}_{\mathbf{c}}(\tau)) = \mathbf{f}(\hat{\mathbf{b}}(\tau + c, \tau), \bar{\mathbf{x}}_{\mathbf{c}}(\tau)) \quad \text{with} \quad \bar{\mathbf{x}}_{\mathbf{c}}(\tau) = \hat{\mathbf{x}}(\tau + c, \tau), \tag{1}$$

which exhibit an information transport along straight lines in the direction of the diagonal. To exploit this, we discretise the DAE (1) on a characteristic grid, which will be described in the following section.

3. Wavelets for Adaptive Grid Generation

We have to determine the MPDAE-solution over the rectangle $[0, T_1] \times [0, T_2]$ of the periodicities, which is depicted in figure 1 (left), and aims at the limit cycle for our quasi-periodic problem. The representation of the domain's diagonal, which contains the solution $\mathbf{x}(t) = \hat{\mathbf{x}}(t, t)$ of the original network equations, is indicated by the dotted lines; the solid lines show the characteristic curves, along which we have to solve system (1), only. The periodicity of the MPDAE-solution $\hat{\mathbf{x}}$ in t_2-direction leads to boundary conditions for the restrictions $\bar{\mathbf{x}}_c$ via interpolation.

To determine the adaptive grids in t_2-direction we integrate the original DAE-system along the characteristic curves $[c, c + T_2]$ in advance using a standard time integration algorithm. Thereby it is crucial to specify consistent initial values for each characteristic curve, which can be obtained by an implicit Euler-step. Then the obtained solutions are decomposed performing discrete wavelet transforms (DWTs). The wavelets used for this transformation are hat-functions, which are "folded" at the edges of the interval $[0, T_2]$ following [4]. In this way a multiresolution analysis of $L^2([0, T_2])$ is constructed and the time-frequency localisation property of the wavelets can be used to generate an adaptive grid by 'simply' inspecting the magnitude of the wavelet coefficients.

For the time-integration along the characteristic curves we do not have to solve for the limit cycle, as we only need the basic structure of the solution to determine an adaptive grid, which could possibly look like the one depicted in figure 1 (right).

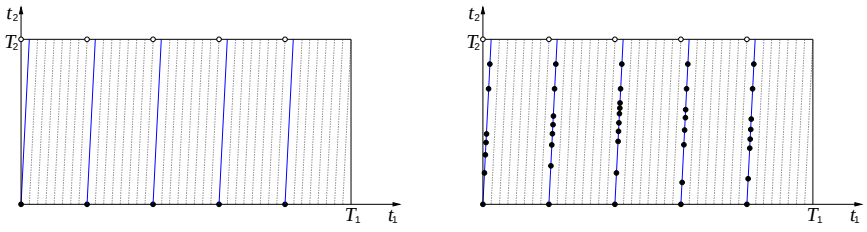

Figure 1. Characteristic curves (left) and adaptive grid (right).

Equipped with a grid tailored to the special structure of the solution, we solve the DAEs (1) using a finite difference discretisation described in [6]. As the equations are only coupled via the interpolation for the boundary conditions, the arising linear system in the Newton iteration is very sparse and can be solved efficiently.

4. Simulation Results and Conclusions

To illustrate the adaptive grid generation presented in this paper, we introduce an industrial test example, the Miller integrator shown in figure 2 (left), which can be described by a set of index-1 differential-algebraic equations. This circuit comprises the two major properties we want to focus on, namely widely separated time scales and heterogeneous signal structures. Apart of a slowly varying harmonic input signal (v_{in}), two pulse functions (p_a and p_b) are involved, which work on a much faster time scale than the input and are characterised by steep gradients.

Figure 2 (middle) shows the adaptive grid on the two-dimensional domain obtained after time-integration along the characteristic curves and DWT of those solutions. Comparing this grid with the MPDAE-solution at node 1 depicted, figure 2 right, we notice that the steep gradients arising due to the pulse functions are perfectly detected by the wavelet transforms.

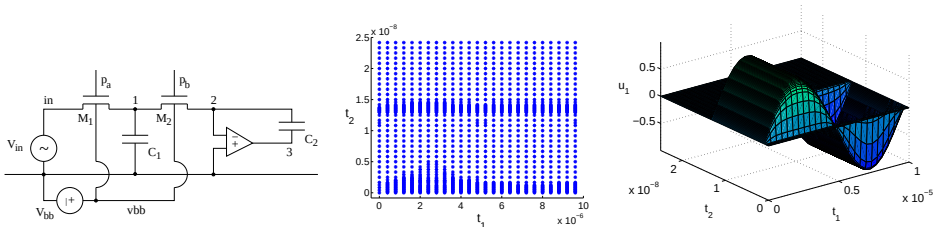

Figure 2. Circuit (left), grid (middle), MPDAE-solution (right).

In conclusion, we have demonstrated on the above example (Miller integrator) that tailored grids can be defined using a pre-simulation in time domain plus DWT with hat-wavelets. This also yields different grids on different characteristic curves. It is Crucial for efficiency that the selection of discretisations by the DWT coefficients is appropriate. Therefore a general algorithm has to be based on a larger set of industrial test examples.

References

[1] G. Alì, A. Bartel, M. Günther, *Parabolic differential-algebraic models in electrical network design,* SIAM J. Mult. Model. Sim., **4**:3 (2005), 813–838.

[2] A. Bartel, M. Günther, *From SOI to abstract electric-thermal-1D multiscale modeling for first order thermal effects*, Math. Comput. Modell. Dynam. Syst., **9**:1 (2003), 25–44.

[3] H.G. Brachtendorf, G. Welsch, R. Laur, A. Bunse-Gerstner, *Numerical steady state analysis of electronic circuits driven by multi-tone signals*, Electrical Engineering **79** (1996) 103–112.

[4] A. Cohen, I. Daubechies, P. Vial, *Wavelets on the interval and fast wavelet transforms*, Appl. Comput. Harmonic Analysis **1** (1993), 54–81.

[5] S. Knorr, M. Günther, *Index Analysis of Multirate Partial Differential-Algebraic Systems in RF-Circuits*, to appear in: Anile, A.M., Alì, G., Mascali, G. (Ed.), Proceedings of the SCEE 2004 conference, Springer-Verlag, Berlin.

[6] R. Pulch, M. Günther, *A method of characteristics for solving multirate partial differential equations in radio frequency application*, Appl. Numer. Math. **42** (2002), 397–409.

Dynamic Optimization Using Control Parameterization

Paul I. Barton

Prof. Barton is the Director of the Process Systems Engineering Laboratory at MIT. He is a chemical engineer by training but also has many interests in applied mathematics and numerical analysis. The broad theme of his research is the modeling, simulation, optimization and design of large-scale dynamic systems encountered in chemical engineering. Applications are drawn from the traditional chemical process industries, and also from less traditional areas such as pharmaceutical and biochemical processes, micro-scale chemical process (e.g., for portable power generation), signaling and regulation networks in biological systems, complex chemical reaction mechanisms such as those in combustion systems, nuclear hydrogen generation, design of OLED displays, and natural gas production, distribution and processing networks. His research interests and contributions include hybrid (discrete/continuous) dynamic systems; design and modeling of complex distillation systems; numerical analysis of ordinary differential, differential-algebraic and partial differential-algebraic equations; sensitivity analysis and automatic differentiation; pollution prevention in process design; mixed-integer and dynamic optimization theory and algorithms; process safety analysis; open process modeling software. Besides these general interests the following paragraphs describe current research efforts relating to the theme of the workshop.

Global Dynamic Optimization. Deterministic global optimization algorithms guarantee an ε–accurate estimate for a global solution of a nonconvex optimization problem in finite computational time. This effort is extending existing notions such as branch-and-bound algorithms to dynamic optimization problems. The optimization problem is formulated on a Euclidean space, and the real valued optimization variables influence the objective and constraint functionals through the solutions of ODEs, DAEs or PDAEs, which are evaluated using numerical integration. Optimal control problems may be addressed within this framework via control parameterization (i.e., approximation of controls in terms of a finite series of basis functions).

Branch-and-bound approaches for global optimization require the construction of convergent *convex relaxations* of the nonconvex functions involved in an optimization problem. A convex relaxation is a convex function that underestimates the nonconvex function on some set of interest. A key question is how to construct, in a computationally tractable manner, tight convergent convex relaxations of functionals with ODEs, DAEs or PDAEs embedded. Our recent research has developed methods for constructing convex relaxations of functionals with linear and nonlinear ODEs embedded. A subsidiary question raised by these methods is how to estimate the image of a subset of a Euclidean space under the solution of ODEs, DAEs or PDAEs. The estimates generated from traditional ideas such as differential inequalities are often too weak for practical application. Our current focus is on exploiting the structure of models of physical systems to tighten

these estimates, new nonlinear convex relaxations, and extensions of these ideas to DAEs and PDAEs.

The ability to construct convex relaxations also enables the application of mixed-integer nonconvex optimization algorithms to solve mixed-integer dynamic optimization (MIDO) problems (dynamic optimization problems that involved a mixture of integer and real valued decisions), which we have demonstrated recently. Looking to the future, approaches for the global solution of semi-infinite and bilevel programs with dynamic systems embedded now seem conceivable.

Large-Scale Dynamic Optimization. This research is investigating methods for finding local solutions of dynamic optimization problems on Euclidean spaces with very many optimization variables (1,000s–100,000s). Usually, the dynamical system will also involve very many state variables (1,000s–1,000,000s), e.g., from a method of lines discretization of a PDAE. Our approach is based on a method for computing Hessian-vector products of ODE embedded functionals for a small multiple of the cost of simulating the ODEs. In particular, this multiple does not change with the number of optimization variables. This ability to compute "cheap" Hessian-vector products enables the application of large-scale optimization methods that do not rely on sparsity, e.g., truncated Newton methods. Extensions of this approach to DAEs are the subject of current research. A potential application for these ideas is to multiple shooting methods for dynamic optimization, which by their nature introduce many optimization variables into the optimization problem solved.

Simulation and Optimization of PDAEs with a Separation of Time Scales. An application related to the start-up of micro-scale chemical processes for electrical power generation highlighted a class of PDAEs in time and one spatial dimension with a natural separation of time scales. The slow variables are lumped, and the fast variables are hyperbolic with all characteristics pointing in the same direction. If the fast variables are approximated as quasi-steady-state (QSS), a natural decomposition occurs in which an adaptive numerical integrator can be used to solve for the spatial profile of the fast variables at each time step for the slow variables. In start-up problems, where shocksand fronts can develop and move around the spatial domain, an adaptive spatial mesh is highly advantageous to the reliability of the simulation, and this reliability is particularly important for optimizations embedding these simulations. Furthermore, preliminary numerical studies indicate that this approach is several orders of magnitude faster than a uniform spatial mesh that yields comparable accuracy (i.e., the spatial discretization error is comparable to the error introduced by the QSS approximation). Efficient implementation of this concept requires the careful application of state-of-the-art sensitivity analysis algorithms. There appears some scope to extend these ideas to situations in which the fast problem is a BVP, and the slow variables are spatially distributed.

Hybrid Systems. Hybrid (discrete/continuous) dynamic systems exhibit both discrete state and continuous state dynamics that are coupled. A popular hybrid

system model is the *hybrid automaton*, in which the discrete state is represented by a finite set of *modes*, and the continuous dynamics while a particular mode is active are described by ODEs or DAEs associated with the active mode. At any point in time one mode is active, but an instantaneous *switch* can also occur to a different mode described by different equations. Switching can occur as a consequence of an explicit control action, or implicitly as a consequence of some condition on the continuous state variables becoming satisfied. Also, at a switch a *jump* may occur in the continuous states, as described by a transition function that maps the final state in one mode to the initial state in the next mode. This model appropriately describes the dynamics of many physical and technological systems of current interest. We have developed many of the key concepts in the theory and algorithms for simulation, sensitivity analysis and optimization of this hybrid automaton model.

Current research is investigating approaches for solving optimization problems with hybrid automata embedded. Two directions are being pursued. One is a global optimization approach based on mixed-integer dynamic optimization, which is primarily aimed at solving problems in formal safety verification of embedded systems. The other direction is local optimization based on nonsmooth optimization techniques, and the computation of associated quantities such as elements of the generalized gradient.

Sensitivity Analysis of Oscillatory Systems. Oscillatory systems are pervasive, for example, in biological systems. Often it is desirable to compute the sensitivity with respect to parameters of characteristics of oscillations such as period, amplitude, phase and derived quantities based on these. However, conventional sensitivity analysis notions do not directly yield this information. We are developing an efficient computational approach based on a BVP formulation that can compute directly all sensitivity information of an oscillating system. Our current approach applies to ODEs, but extensions to DAEs and hybrid systems would be desirable.

Model Reduction for Chemical and Biological Networks. Modern experimental techniques in conjunction with quantum computational chemistry are facilitating the construction of detailed chemical kinetic models that can predict accurately the formation and destruction of byproducts and pollutants in processes such as combustion, pyrolysis, and super-critical water oxidation. Similarly, high throughput experimental techniques, amongst others, are facilitating elucidation of the biochemical networks that govern phenomena such as signaling and regulation in cells. Today these detailed ODE/DAE models can often involve 100s-1,000s chemical species and 1,000s-10,000s chemical reactions.

It is often desirable to embed these chemical/biochemical kinetic models in reacting flow simulations where it is necessary to repeat the chemistry model at a large number of spatial grid points associated with the semi-discretization of PDAEs, or in cell ensemble or population balance simulations of large groups of cells. The need to repeat the large-scale chemical kinetic model at every spatial

grid point or for each cell in an ensemble or having each species concentration as an independent variable in a population balance can easily overwhelm state-of-the-art algorithms running on advanced computing architectures. It is therefore necessary to consider approaches for reducing the size of the chemical kinetic model while still guaranteeing accuracy in the simulation results. We are developing a fully automated local kinetic model reduction procedure that deletes reactions and/or species from the model using a mixed-integer optimization formulation. This procedure can guarantee finding the smallest possible kinetic model that satisfies user specified error tolerances. Moreover, the resulting reduced model can still be interpreted physically as a subset of the original chemical mechanism. Current research is showing that these notions can be extended to generate models that have rigorous regions of validity, i.e., the prediction of the reduced model is guaranteed to be within some tolerance of the full model for some region of state space. In conjunction with Profs. Green and Tidor at MIT we are also exploring the use of libraries of these reduced models in adaptive chemistry and adaptive biology simulations that adapt the reduced kinetic model to local conditions in order to maintain model accuracy while reducing computational time.

Software. Our goal is always to develop software implementing our research ideas that can be effectively used by a broader community. Jacobian is a modeling environment for hybrid DAE based models that supports model analysis, simulation, sensitivity analysis, parameter estimation and optimization. DAEPACK is an automatic differentiation tool and numerical library implementing many of our symbolic and numeric algorithms. GDOC implements the global dynamic optimization ideas. RIOT implements the model reduction ideas. All of this software is distributed for free for academic use.

Numerical Methods for Efficient Nonlinear Model Predictive Control and Moving Horizon State Estimation

Hans Georg Bock

(joint work with Jan Albersmeyer, Moritz Diehl, Ekaterina Kostina, Peter Kühl, Andreas Schäfer, Johannes P. Schlöder, Leonard Wirsching, Frank Allgöwer, Rolf Findeisen)

The presentation reports on recent progress in the development of numerical methods for the real-time computation of constrained closed-loop optimal controls, and in particular the case of nonlinear model predictive control (NMPC) and moving horizon estimation of states and parameters (MHE), for processes governed by large systems of Differential Algebraic Equation (DAE) as they arise e.g. from semi-discretization of instationary Partial Differential Equations.

Closed-loop optimal controls as in NMPC are important in dynamic processes with uncertainties in order to cope with perturbations of states and systems parameters. One possible way to compute them is to solve the Hamilton-Jacobi-Bellman equation and tabulate the resulting closed-loop solution, which is a cumbersome task in high dimensional state spaces.

An alternative way, followed here, is to solve the corresponding open-loop constrained optimal control problem in real time for the perturbed data, and use the first optimal control instant as the feedback control, see [1, 8]. The challenge here, however, is that in order to be feasible in practice, numerical solution algorithms have to be developed that *minimize the response time with respect to perturbations,* while at the same time dealing with non-linear dynamics and boundary conditions as well as non-linear control and state inequality constraints. For time critical applications, the solution of such problems by choosing even the fastest off-line optimization methods is therefore out of the question, and new methods have to be developed [3, 4].

Of particular interest in the present talk are problems in which the required response times may be orders of magnitude shorter than the time for solving an off-line optimal control problem. As the basic solution approach we choose the *direct multiple shooting method* which is an "all-at-once" optimization method that consists of a finite dimensional parameterization of the control functions and a time discretization - or more precisely a re-parameterization - of the differential-algebraic equations. The result is a transcription to a large-scale constrained non-linear programming problem with a system of non-linear equality constraints that exhibit a special boundary value problem structure, for which very effective solution methods have already been developed.

The direct multiple shooting approach has several advantages that are important in the NMPC and the real-time optimization context: The incorporation of the state variables as unknowns in the optimization process reduces the nonlinearity and improves both local and global convergence, the algorithm allows a numerically *stable treatment of optimal control problems with highly unstable and even chaotic dynamics,* the decomposition of the integration process into independent subintervals allows a convenient parallelization of the computationally intensive parts. Moreover, since in multiple shooting modified standard DAE integrators can be applied, an effective *adaptive discretization error control* of trajectories and derivatives is possible. Starting from here, the so-called "real-time iteration" approach [4] is developed, which integrates, among others, the following algorithmic components:

- a perturbation embedding, which makes the computation of function values and derivatives largely independent of the actual value of system and parameter estimates,
- approximate Newton, Gauss-Newton and quasi-Newton optimization methods for the equality and inequality constrained non-linear programming problem, and

- structure exploiting linear algebra techniques to decompose and solve the quadratic program (QP) sub-problems.

One of its basic features is that now in each iteration of the optimization process, the latest available process data are being used. A reformulation of the optimization algorithm and a pre-computation - as far as possible - of constraint residuals, gradients, Hessians and QP decompositions splits each iteration into a preparation and a feedback phase. As a consequence, the *response times to perturbations of states and systems parameters* are minimized. In real experiments for a high purity distillation column, the response times realized by these new methods have been shown to be orders of magnitude faster than those of previous approaches based on on-line application of off-line optimization methods [6].

For the class of NMPC approaches that guarantee globally stable closed-loop controls, the new approximate scheme could be *proven to be nominally stable* by combining contractivity properties of the optimization algorithm with stability properties of the NMPC scheme [5].

It is further shown that the approach can be drastically accelerated by *special algorithmic schemes for on-line feasibility and optimality improvement,* that utilize the principles of algorithmic and internal differentiation, or quasi-Newton update techniques, for the efficient computation of Jacobians and Lagrange gradients at different accuracy levels [1, 7].

Practical applications of NMPC require a simultaneous state and parameter estimation and computation of the closed-loop NMPC control in real-time. It is shown how the principles of the real-time iteration approach can be extended to the dual problem of a *moving horizon estimation of parameter and states*, and how to effectively nest the real-time iteration for both MHE and NMPC. In order to take the statistical error of the system and parameter estimates and potentially other errors in the model - into account, a *robust NMPC approach* is developed that is based on a computationally efficient worst case optimization algorithm [2, 9].

In addition, we report on specific *reduced* Newton and Gauss-Newton real-time strategies. At the expense of slightly increased response times they minimize the effort for computing derivatives, for problems with a large number of state variables, but comparatively few actual degrees of freedom for the controls, as e.g. in Method of Lines approaches.

References

[1] H.G. Bock, M. Diehl, E.A. Kostina, and J.P. Schlöder, *Constrained optimal feedback control of systems governed by large differential algebraic equations*, in: L. Biegler, O. Ghattas, M. Heinkenschloss, D. Keyes, and B. van Bloemen Waanders (eds), Real-Time and Online PDE-Constrained Optimization, SIAM, 2005.

[2] M. Diehl, H. G. Bock, E. A. Kostina, *An approximation technique for robust nonlinear optimization*, Mathematical Programming, **107** (2006), 213-230.

[3] M. Diehl, H.G. Bock, J.P. Schlöder, R. Findeisen, Z. Nagy, and F. Allgöwer, *Real-time optimization and nonlinear model predictive control of processes governed by differential-algebraic equations*, J. Proc. Contr., **12(4)** (2002), 577-585.

[4] M. Diehl, H.G. Bock, and J.P. Schlöder, *A real-time iteration scheme for nonlinear optimization in optimal feedback control*, SIAM J. Control Optim., **43(5)** (2005), 1714-1736.

[5] M. Diehl, R. Findeisen, F. Allgöwer, H.G. Bock, and J.P. Schlöder, *Nominal stability of the real-time iteration scheme for nonlinear model predictive control* , IEE Proc.-Control Theory Appl., **152(3)** (2005), 296-308.

[6] M. Diehl, R. Findeisen, S. Schwarzkopf, I. Uslu, F. Allgöwer, H.G. Bock, E.D. Gilles, and J.P. Schlöder, *An efficient algorithm for nonlinear model predictive control of large-scale systems. Part II : Application to a distillation column*, Automatisierungstechnik , **51(1)** (2003), 22-29.

[7] M. Diehl, A. Walther, H.G. Bock, and E.A. Kostina, *An adjoint-based SQP algorithm with quasi- Newton Jacobian updates for inequality constrained optimization*, Technical Report MATH-WR-02-2005, Technical University Dresden, 2005.

[8] R. Findeisen, M. Diehl, T. Bürner, F. Allgöwer, H.G. Bock, and J.P.Schlöder, *Efficient output feedback nonlinear model predictive control*, in Proc. Amer. Contr. Conf., Anchorage (2002), 4752-4757.

[9] P. Kühl, A. Milewska, M. Diehl, E. Molga and H. G. Bock, *NMPC for runaway-safe fed-batch reactors*, in Proc. Int. Workshop on Assesment and Future Directions of NMPC (2005), 467 - 474.

Some Aspects of the Optimal Control of Nonlinear Differential-Algebraic Equations

Rainer Callies

Conditions for the existence of Lagrangian multiplier functions in nonlinear optimal control problems for DAEs are derived via a modified embedding technique. We consider the following prototype of a variational problem for DAEs

$$\left\{\begin{array}{l} J(y) := \displaystyle\int_{t_a}^{t_f} f(t,y,\dot{y})\,\mathrm{d}t \stackrel{!}{=} \min \\ y(t) := (y_1(t),\ldots,y_n(t))^T\,,\ \ y \in C_c^1([t_a,t_f],\mathbb{R}^n) \\ y(t_a) = y_a\,,\ \ y(t_f) = y_f \\ h(t,y) = 0 \in \mathbb{R}\ \ \wedge\ \ \|\nabla h(t,y_0)\|_2 > 0\,,\ \nabla := (\partial/\partial y_1,\ldots,\partial/\partial y_n)^T \end{array}\right.$$

with the optimal solution $y_0(t)$.

For an elegant proof of the necessary conditions an embedding formula is chosen which projects any function $\eta(t)$ on the algebraic manifold

$$\begin{aligned} y(t,\varepsilon) &:= y_0(t) + \varepsilon\tilde{\eta}(t) + \alpha(t,\varepsilon)\cdot\nabla h(t,y_0(t))) \\ \tilde{\eta}(t) &:= \eta(t) - \left[\nabla h^T(t,y_0(t))\cdot\eta(t)\right]\frac{\nabla h(t,y_0(t))}{\|\nabla h(t,y_0(t))\|_2^2} \end{aligned}$$

with $\eta \in C_c^1([t_a,t_f],\mathbb{R}^n)\ \ \wedge\ \ \eta(t_a)=0$, $\eta(t_f)=0$. $\alpha(t,\varepsilon)$ is determined such that $\forall\, t \in [t_a,t_f]$, $\forall\, \varepsilon \in [-\varepsilon_0,\varepsilon_0]$

$$h(t,y_0(t) + \varepsilon\tilde{\eta}(t) + \alpha(t,\varepsilon)\cdot\nabla h(t,y_0(t))) \equiv h(t,y_0(t))\,.$$

It is shown in the talk, that $\alpha(t,\varepsilon) \in \mathbb{R}$ exists, is uniquely determined and contains all the nonlinear and no linear portions of the embedding. This embedding formula separates the linear from the nonlinear parts of the embedding and fulfils the

algebraic constraint: $\varepsilon = 0$ is an interior point. The first necessary condition reads as

$$\begin{aligned} 0 &= \frac{\mathrm{d}J}{\mathrm{d}\varepsilon}(\varepsilon)\bigg|_{\varepsilon=0} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left(\int_{t_a}^{t_f} f(t, y(t,\varepsilon), \dot{y}(t,\varepsilon))\,\mathrm{d}t\right)\bigg|_{\varepsilon=0} \\ &= \int_{t_a}^{t_f} \sum_{i=1}^{n} f_{y_i}(t, y_0, \dot{y}_0)\left(\eta_i - \frac{\partial h}{\partial y_i}(t, y_0)\cdot\left\{\frac{\left[\nabla h^T(t, y_0)\cdot\eta\right]}{\|\nabla h(t, y_0)\|_2^2}\right\}\right)\mathrm{d}t \\ &+ \int_{t_a}^{t_f} \sum_{i=1}^{n} f_{y_i'}(t, y_0, \dot{y}_0)\left(\eta_i' - \left\{\frac{\partial h}{\partial y_i}(t, y_0)\cdot\frac{\left[\nabla h^T(t, y_0)\cdot\eta\right]}{\|\nabla h(t, y_0)\|_2^2}\right\}'\right)\mathrm{d}t \end{aligned}$$

Integration by parts together with the definition of the multiplier function

$$\mu(t) := \sum_{i=1}^{n}\left(f_{y_i}(t, y_0, \dot{y}_0) - \frac{\mathrm{d}}{\mathrm{d}t}f_{y_i'}(t, y_0, \dot{y}_0)\right)\frac{\partial h}{\partial y_i}(t, y_0)\frac{1}{\|\nabla h(t, y_0)\|_2^2}$$

yields the well-known Euler-Lagrange equations

$$F_{y_i} - \frac{\mathrm{d}}{\mathrm{d}t}F_{y_i'} = 0\,, \quad F := f + \mu h\,, \quad i = 1, \ldots, n\,.$$

The last step and thus a reasonable definition of the multiplier function is possible only, if the respective terms in the first variation do not vanish completely.

An example of the latter situation is given in [1]:

Let us find $y : [0, T] \to \mathbb{R}^4$, which minimizes

$$J(y) := \frac{1}{2}y_1^2(T) + \frac{1}{2}\int_0^T \left(y_3^2(t) + u^2(t)\right)\,\mathrm{d}t\,, \quad y := (y_1, y_2, y_3, u)^T$$

subject to the DAE conditions (DAE index 2) and boundary conditions

$$\begin{aligned} \dot{y}_1(t) &= u(t) & \qquad y_1(0) &= y_{10} \\ \dot{y}_2(t) &= -y_3(t) + u(t) & \qquad y_2(0) &= 0 \\ 0 &= y_2(t) = h(t, y(t)) \quad \forall\, t \in [0, T] \end{aligned}$$

Insertion into the first variation yields

$$0 = \int_{t_a}^{t_f} \sum_{\substack{i=1 \\ i\neq 2}}^{n} \left(f_{y_i}(t, y_0, \dot{y}_0)\eta_i(t) + f_{y_i'}(t, y_0, \dot{y}_0)\eta_i'(t)\right)\,\mathrm{d}t\ .$$

The first variation contains no information about the DAE constraint, no information about the η_2-component and no Lagrangian multiplier function; the standard approach via the extended Hamiltonian fails. The existence of a Lagrangian multiplier function – and thus of the classical DAE boundary value problem – is not essential for a well-defined optimal control problem. In those cases, the (partial) transformation to minimum coordinates is a more powerful approach.

As an example from industrial applications, the time-optimal motion of a three-link manipulator is investigated with its end-effector following a prescribed path in space. The dynamic equations together with the spatial restrictions form a nonlinear differential-algebraic system of differential index 3. Additionally, the

optimal control problem for this DAE system contains multiple restrictions and several interior points. A partial transformation to minimum coordinates is an elegant way to eliminate severe mathematical problems of the type discussed above, which arise from the special structure of the algebraic constraints [2]. The transformation results in a system of linear equations of motion. Introducing nonlinear control/state constraints is the price to be paid for the much simpler structure of the overall problem. The transformed optimal control problem is transferred into a nonlinear multi-point boundary value problem and solved numerically by the advanced multiple shooting method JANUS. The switching structure is derived automatically. By backward transformation the actuator torques and their switching behaviour are easily obtained. Solutions are presented for which the set of feasible controls reduces to a single point at certain times.

References

[1] A. Backes, *A necessary optimality condition for the linear-quadratic DAE control problem*, Institut für Mathematik, HU Berlin, Preprint 16, 2003.

[2] S. Breun, R. Callies, *Redundant Optimal Control of Manipulators Along Specified Paths*, III Europ. Conference on Computational Mechanics, Lisbon, Portugal, 2006.

What's New with Direct Transcription Methods for Optimal Control Problems?

Stephen L. Campbell

(joint work with J. T. Betts, Anna Engelsone)

Direct transcription methods [1] have been successfully used to solve a variety of optimal control problems for at least 20 years. They proceed by totally discretizing the optimal control problem and then passing the fully discretized problem, which is now a nonlinear programming problem (NLP), to a NLP solver. Direct transcription methods are used in a number of industrial codes such as SOCS (Sparse Optimal Control Software) developed at Boeing. A number of papers have been written about direct transcription methods. The computational examples in this talk were solved using SOCS but our comments are applicable to other optimal control solvers with a similar philosophy.

The dynamical constraints in optimal control problems are often differential equations. Accordingly the analytic and numerical theory for differential equations plays an important role in optimal control theory and in the analysis and design of algorithms.

Sometimes the dynamical constraints are Differential Algebraic Equations (DAEs) rather than ordinary differential equations. In practice most optimal control problems have various position and velocity constraints. When these constraints are active we again have a DAE. This has motivated the application of DAE theory to the examination of constrained optimal control problems.

Given that so much work has been done on DAEs and also on direct transcription methods for optimal control problems, it is natural to ask whether anything

remains to be done when studying such methods. In recent years it has begun to be realized that the existing computational theory needs to be modified when talking about direct transcription methods applied to constrained optimal control problems. Also, some modifications may be necessary if the approach is going to be extended to an even wider class of problems in order to encompass more applications. In this talk we present two illustrations from our current investigations. These two examples are: use of the theory for constrained differential equations and their implications for providing guidance to users of direct transcription software and the estimation of adjoint variables. There are other recent examples [3, 2]. We focus here on only the most recent research [4, 5, 6, 7].

Traditionally a controlled system is viewed as a differential equation in the state variable with an input function which is the control. The control is considered as a forcing function. The numerical and analytical theory of such a differential equation is then applied. If there are no control constraints, then the usual advice in forming a regularized problem is to make sure the control is weighted in the cost. If constraints are present and the resulting problem is solved by either control parameterization or solution of the necessary conditions, then the DAE theory is applied in a similar way. However, with direct transcription the situation is different since the optimization software treats the algebraic part of the state and what the user considers the control in the same way. In this talk it is shown that what is important is that there is some alternative choice of control (which a user does not need to find) for which this "virtual control" leads to an index one system. Furthermore, what is important for obtaining a numerical solution is that this virtual index one control is positively weighted in the cost. Thus the way to regularize when a DAE is present is to include all the algebraic variables in the cost.

One advantage of direct transcription is that it is not necessary to compute the adjoint variables from the necessary conditions for the optimal control problem. However, sometimes estimates of the adjoint variables are useful or needed. This can happen when computing sensitivities. There are also results which state that under certain conditions the multipliers from the discrete approximation provide estimates of the adjoint variables which are more accurate than the control estimates [8]. These improved adjoint estimates can then be used, in the case of unconstrained problems, to give improved control estimates.

The two primary discretization methods used by SOCS are the trapezoid (TR), which is second order, and Hermite Simpson (HS), which is fourth order as an integrator. Computational studies show that previous theoretical results are not correctly predicting the values of the adjoint estimates. A careful analysis proves that this is due to the fact that the HS and the TR are implemented in a compressed form for computational efficiency rather than in the Butcher array formulation which is the usual theoretical starting point. Furthermore, it is shown in this talk how to modify how the multipliers are used and get improved estimates. In the case of TR, the Butcher array based estimates are then obtained. In the case of HS, second order is obtained.

References

[1] J. T. Betts, *Practical Methods for Optimal Control using Nonlinear Programming*, SIAM, Philadelphia, 2001.

[2] J. T. Betts, S. L. Campbell, and A. Engelsone, *Direct transcription solution of optimal control problems with higher order state constraints: theory vs practice*, Optimization and Engineering, to appear.

[3] J. Betts, N. Biehn, and S. L. Campbell, *Convergence of nonconvergent IRK discretizations of optimal control problems with state inequality constraints*, SIAM J. Sci. Comp., **23** (2002), 1981-2007.

[4] S. L. Campbell and R. Marz, *Direct transcription solution of high index optimal control problems and regular Euler-Lagrange equations*, J. Comp. Appl. Math., to appear.

[5] A. Engelsone, S. L. Campbell, and J. T. Betts, *Order of Convergence in the Direct Transcription Solution of Optimal Control Problems*, Proc. IEEE Conf. Decision Control - European Control Conference, Seville, Spain, 2005.

[6] A. Engelsone, S. L. Campbell, and J. T. Betts, *Direct Transcription Solution of Higher-Index Optimal Control Problems and the Virtual Index*, Appl. Numerical Mathematics, to appear.

[7] A. Engelsone and S. L. Campbell, *Adjoint Estimation using Direct Transcription Multipliers: Trapezoidal Method*, Preprint, 2006.

[8] W.W. Hager, *Runge-Kutta Methods in Optimal Control and the Transformed Adjoint System*, Numerische Mathematik **87** (2000) 247–282.

Exponential integrators for convection dominated flows

Elena Celledoni

We consider nonlinear convection diffusion problems with a dominating convection term. These models are challenging and ubiquitous in applications, an example being the numerical simulation of internal waves phenomena occurring between the layers of a stratified flow. In Norwegian fjords, layers of stratified water with different temperature and salt concentration occur due to ice melting and freshwater supply from rivers. Internal waves are caused by the tide and have a dramatic influence on the ecosystem. The Navie-Stokes equations with the Bousinessque approximation are a popular tool for modelling these phenomena. We consider convection diffusion PDE models depending on a viscosity parameter ν of the type,

$$\frac{\partial}{\partial t}u(\mathbf{x},t) + \mathbf{V}\cdot\nabla u(\mathbf{x},t) = \nu\nabla^2 u + f(\mathbf{x}), \tag{1}$$

with $\mathbf{x}\in\Omega\subset\mathbf{R}^d$ and $\mathbf{V}:\mathbf{R}^d\times[0,T]\to\mathbf{R}$ is a given vector field, but can also be $V=u$, $u:\mathbf{R}^d\times[0,T]\to\mathbf{R}^d$, and $u(\mathbf{x},0)=u_0(\mathbf{x})$. The case when the parameter ν tends to zero is particularly interesting and very challenging from the numerical point of view. In this case the numerical discretizations often lead to phenomena of numerical dispersion. A suitable generalization of these equations leads to the Navier–Stokes equations at the presence of high Reynold's numbers.

We present a new class of integration methods which present good performance for convection dominated problems. These methods are exponential integrators of Runge-Kutta type. They allow for the solution of just one linear system per

stage, and their advantage is that they allow for the very accurate integration of the linearized convection. As the diffusion is treated implicitly and the convection is accurately resolved, the methods have superior properties of linear stability.

A simple example is the following transport-diffusion algorithm studied in [5],

$$\begin{aligned} &\frac{Du_{n+\frac{1}{2}}}{Dt} = 0, \;\; u_{n+\frac{1}{2}}(x,t_n) = u_n(x), \;\; \text{on } [t_n, t_n+h] \\ &u_{n+\frac{1}{2}}(x) = u_{n+\frac{1}{2}}(x, t_n+h) \\ &u_{n+1} = u_{n+\frac{1}{2}} + h\nu\nabla^2 u_{n+1} + hf, \end{aligned} \tag{2}$$

the convecting vector field is $\mathbf{V}(x) = u_n(x)$. The integration of a linear pure convection problem arises as a building block of the integration method and can be achieved by computing characteristics, as follows

$$\begin{aligned} &u_{n+\frac{1}{2}}(x) = u_{n+\frac{1}{2}}(x, t_n+h) = u_n(X(t_n)) \\ &\frac{dX}{d\tau} = u_n(X(\tau)), \;\; X(t_n+h) = x. \end{aligned} \tag{3}$$

The equation for the characteristics $X(\tau)$ must by integrated backwards in time, either exactly or with a suitable numerical integrator.

The study of higher order time integrators of this type is also motivated by some observations in some recent work by Karniadakis et al., [6]. The authors point out that the use of low order semi-implicit methods in the case of direct numerical simulation of turbulent flows leads to prohibitive time step restrictions. In fact the time-step dictated by the CFL condition can be of several orders of magnitude smaller then the intrinsic temporal scale of the problem predicted by the theory. Exponential Runge-Kutta integrators can overcome this time step restriction.

Preliminary work illustrating the potential of the methods has been presented in [2]. This work is also related to the methods presented in [3], for the disretization of the Navier-Stokes equations.

Extending the methods to the case of incompressible Navier-Stokes equations while maintaining high order requires taking into account the classical theory of DAEs along the lines of recent work by Petzold and Zheng [4].

References

[1] E. Celledoni, A. Marthinsen and B. Owren, *Commutator-free Lie group methods*, FCGS, **19** (2003), 341–352.

[2] E. Celledoni, *Eulerian and semi-Lagrangian commutator-free exponential integrators*, CRM Proceedings and Lecture Notes, **39** (2005), 77–90.

[3] Y. Maday, A. T. Patera and E. M. Rønquist, *An operator integration factor splitting method for time dependent problems: Application to incompressible fluid flows*, J. of Sci. Comp., **5** (1990), 263–292.

[4] L. Petzold and Z. Zheng, *Runge-Kutta-Chebyshev projection method*, Report of the Department of UC Santa Barbara, CA 93106, USA (to appear).

[5] O. Pirroneau, *On the transport-diffusion algorithm and its applications to the Navier-Stokes equations* , Numer. Math. 38, 309-332 (1982).

[6] D. Xiu and G.E. Karniadakis, *A semi-Lagrangian high-order method for Navier–Stokes equations*, J. of Comput. Phys. **172** (2001), 658–684.

A quadratic regulator problem related to singular systems in Hilbert space

Angelo Favini
(joint work with Luciano Pandolfi)

Of concern is the quadratic regulator problem for a plant described by

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Gu,\ x(0) = x_0 \in X, \end{aligned}$$

where A, B and G are operators with the following properties:

- A generates an analytic semigroup on a Hilbert space X;
- the operators B and G are linear and continuous from U into X and from U into Y, where U and Y are Hilbert spaces;
- $\mathrm{Ker}(G^*) = 0$.

The cost functional to minimize is

$$J(x_0; u) = \int_0^T F(x(t), u(t))dt + \langle \begin{bmatrix} x(T) \\ y(T) \end{bmatrix}, M \begin{bmatrix} x(T) \\ y(T) \end{bmatrix} \rangle,$$

where

$$F(x, u) = \langle \begin{bmatrix} x \\ y \end{bmatrix}, Q \begin{bmatrix} x \\ y \end{bmatrix} \rangle + |u|^2,$$

Q, M are symmetric nonnegative continuous linear operators and

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{12}^* & M_{22} \end{bmatrix}, \quad M_{22} > cI > 0.$$

Even if the optimal control in general does not exist, it is seen that, when it exists, it admits a variational characterization. A necessary and sufficient condition for existence of an optimal control is described by means of a compatibility relation. We are going to study this quadratic regulator problem with $u \in L^2(0, T; U)$. Clearly, the quadratic cost is not defined for every $u \in L^2(0, T; U)$. So, we introduce a suitable domain over which the quadratic cost makes sense.

Let

$$\operatorname{ess} \lim_{t \to T-} u(t) = l$$

when for every $\epsilon > 0$ there exists $\delta > 0$ such that the following set has zero Lebesgue measure:

$$\{t\,,\ \text{such that } ||u(t) - l|| > \epsilon\,,\ 0 < T - t < \delta\}\,.$$

If two functions u and u' belongs to the same equivalent class $[u] \in L^2(0, T)$ then the ess lim exists for one of them if it exists for the second, and the limit itself is the same.

We introduce the linear space $\mathcal{U}$ of those equivalent classes in $L^2(0, T; U)$ identified by a representative u such that the essential limit for $t \to T-$ exists and we define

$$u(T) = \operatorname{ess} \lim_{t \to T-} u(t).$$

The subspace $\mathcal{U} \subseteq L^2(0,T;U)$ just defined is the domain over which the problem will be studied. Its introduction is suggested in [1].

Due to the fact that $u(\cdot) \to J(x_0;u(\cdot))$ is neither continuous nor closed, we must check explicitly that when the optimal control exists, it has a variational characterization. We proceed as usual: if the optimal control $u^+(\cdot;x_0)$ exists, then we can compute $J(x_0;u^+(\cdot;x_0)+v(\cdot))$, $v(\cdot)\in\mathcal{U}$. By taking $v(\cdot)$ "concentrated" near the final point T we can separate the distributed and terminal conditions and we find that, see [2], [3] for more details,

$$\begin{aligned}
&x(0)=x_0\\
&\dot{x}^+(t;x_0)=Ax^+(t;x_0)+Bu^+(t;x_0)\\
&u^+(t;x_0)=-B^*p(t)-G^*q(t)\\
&\dot{p}=-A^*p-Q_{11}x^+(t;x_0)-Q_{12}y^+(t;x_0)\\
&p(T)=M_{11}x^+(T;x_0)+M_{12}y^+(T;x_0)\\
&q(t)=Q_{12}^*x^+(t;x_0)+Q_{22}y^+(t;x_0)\\
&y^+(t;x_0)=Gu^+(t;x_0)\,.
\end{aligned}$$

Moreover, the following "consistency condition" must hold:

$$M_{22}Gu^+(T;x_0)=-M_{12}^*x^+(T;x_0)\,.$$

Conversely, it is easily seen that if a pair $x(t;x_0,u)$ and $u(t)$ is related in this way, then for every $v\in\mathcal{U}$ the cost with $u(t)+v(t)$ is the sum of two squares, and it is minimum for $v=0$. Hence:

Theorem 1 Let $x_0\in X$ and let us consider the following two-point problem:

$$(1)\qquad \begin{cases}
x(0)=x_0\\
\dot{x}=\left\{A-B(I+G^*Q_{22}G)^{-1}G^*Q_{12}^*\right\}x(t)\\
-B(I+G^*Q_{22}G)^{-1}B^*p(t)\\
\dot{p}=-\left[A-B(I+G^*Q_{22}G)^{-1}G^*Q_{12}^*\right]^*p(t)\\
-\left\{Q_{11}-Q_{12}G(I+G^*Q_{22}G)^{-1}G^*Q_{12}^*\right\}x(t)\\
p(T)=\left\{M_{11}-M_{12}M_{22}^{-1}M_{12}^*\right\}x(T)\,.
\end{cases}$$

If the optimal control $u^+(\cdot;x_0)$ exists then this two-point problem is solvable and the optimal control satisfies the "compatibility condition" (3). Conversely, let the two-point problem (1) be solvable and let us define

$$(2)\qquad u^+(t;x_0)=-(I+G^*Q_{22}G)^{-1}B^*\left\{p(t)+G^*Q_{12}^*x(t)\right\}\,.$$

The optimal control for the initial condition x_0 exists (and then it is given by (2)) if and only if the following "compatibility condition" is satisfied:

$$(3)\qquad Gu^+(T;x_0)=-M_{22}^{-1}M_{12}^*x(T)\,.$$

Proof. The direct implication was proved already. Conversely, let the two-point problem be solvable and let u be defined as in (2). It is easy to see that all the conditions for the optimal control are satisfied, provided that the compatibility condition (3) holds. ∎

Now we note that the two-point problem (without the compatibility condition) is always solvable:

Theorem 2 We have

$$M_{11} - M_{12}M_{22}^{-1}M_{12}^* \geq 0 \tag{4}$$

so that the two-point problem (1) is solvable. Moreover, the vector $x(T)$ is a linear and continuous function of x_0.

Proof. The positivity condition (4) follows from $M \geq 0$.

The two-point problem (without the compatibility condition) can be written as follows:

$$\begin{aligned} x(0) &= x_0, \\ \dot{x} &= \mathcal{A}x + Bu\,, \qquad \mathcal{A} = A - B(I + G^*Q_{22}G)^{-1}G^*Q_{12}^*\,, \\ \dot{p} &= -\mathcal{A}^*p - \mathcal{Q}x\,, \qquad \mathcal{Q} = Q_{11} - Q_{12}G(I + G^*Q_{22}G)^{-1}G^*Q_{12}^*\,, \\ p(T) &= \mathcal{M}x(T), \qquad \mathcal{M} = M_{11} - M_{12}M_{22}^{-1}M_{12}^*\,. \end{aligned}$$

This is the two-point problem of the quadratic cost

$$\tilde{J}(x_0;u) = \int_0^T \Big\{ \langle \mathcal{Q}x(t), x(t)\rangle + \langle \mathcal{R}u(t), u(t)\rangle \Big\}\, dt + \langle \mathcal{M}x(T), x(T)\rangle$$

where $\mathcal{R} = (I + G^*Q_{22}G)$ is coercive and $\mathcal{M} \geq 0$ see condition (4). Hence there exists a unique optimal control for every x_0. Furthermore, if $\tilde{x}(t;x_0)$ is the optimal trajectory of x_0, then the transformation $x_0 \to \tilde{x}(t;x_0)$ is linear and continuous. ∎

We use the last statement of the theorem as follows: the compatibility condition (3) can be written as

$$\begin{aligned} &\mathcal{L}x(T) = 0 \\ &\mathcal{L} = G(I + G^*Q_{22}G)^{-1}B^*\left\{M_{11} - M_{12}M_{22}^{-1}M_{12}^* + G^*Q_{12}^*\right\} + M_{22}^{-1}M_{12}^*\,. \end{aligned} \tag{5}$$

Hence:

Theorem 3 The initial condition x_0 is optimizable if and only if the component $x(T)$ of the solution of the two-point problem belongs to $\ker \mathcal{L}$. In particular:

- the set of the optimizable initial conditions is a closed subspace of X;
- every initial condition x_0 is optimizable if and only if $\mathcal{L} = 0$.

Of course, the optimal control $\tilde{u}(t;x_0)$ for $\tilde{J}$ is given by

$$\tilde{u}(\cdot;x_0) = -\mathcal{R}^{-1}B^*p(\cdot)\,.$$

It always exists while $u^+(\cdot;x_0)$ exists only if x_0 is optimazable. The previous considerations show:

Theorem 4 If x_0 is optimizable then $\tilde{u}(\cdot;x_0) = u^+(\cdot;x_0)$.

The Riccati equation is an important tool in the quadratic regulator problem. Hence we now relate our problem to a differential Riccati equation. It is known that the component p of the two-point problem (1) is expressed as

$$p(t) = \mathcal{P}(t)x(t)$$

where $\mathcal{P}(t)$ solves the Riccati equation

$$\frac{d}{dt}\langle \mathcal{P}(t)x, y\rangle = -\langle \mathcal{A}x, \mathcal{P}(t)y\rangle - \langle \mathcal{P}(t)x, \mathcal{A}y\rangle - \langle \mathcal{Q}x, y\rangle$$
$$+\langle B^*\mathcal{P}(t)x, \mathcal{R}^{-1}B^*\mathcal{P}(t)y\rangle\,, \qquad \forall x\,,\ y \in \mathrm{dom}\mathcal{A}\,; \qquad \mathcal{P}(T) = \mathcal{M}$$

so that the optimal control of $\tilde{J}(x_0; u)$ is

$$\tilde{u}(t) = \mathcal{R}^{-1}B^*\mathcal{P}(t)x(t)$$

where now x solves the closed loop equation

$$\dot{x} = \Big[\mathcal{A} - B\mathcal{R}^{-1}B^*\mathcal{P}(t)\Big]x\,, \qquad x(0) = x_0\,. \tag{6}$$

As we noted, this is also the optimal control of $J(x_0; u)$ when x_0 is optimizable. Hence,

Theorem 5 Let x_0 be optimizable. Then, the optimal control has the feedback form

$$u^+(t; x_0) = \mathcal{R}^{-1}B^*\mathcal{P}(t)x^+(t; x_0)\,.$$

References

[1] Clements D.J., Anderson B.D.O., Singular optimal control: the linear-quadratic problem, Lecture Notes in Control and Information Sciences, Vol. 5. Springer-Verlag, Berlin-New York, 1978.

[2] Favini A., Pandolfi L., A quadratic regulator problem for singular systems, to appear.

[3] Favini A., Pandolfi L., A quadratic regulator problem related to identification problems and singular systems, submitted.

Towards Explicit Methods for DAEs

C. William Gear

Explicit methods have previously been proposed for parabolic PDEs and for stiff ODEs. RKC methods [2, 3] handle problems with eigenvalues distributed along the negative real axis, while projective methods [1] can be designed to handle groups of widely separated eigenvalues. We discuss ways in which Differential Algebraic Equations might be regularized so that they can be efficiently integrated by explicit projective methods and illustrate the effectiveness of this approach for some simple index three problems. The approach places very stiff regularizing eigenvalues at known locations and damps their components in one or two inner steps of the projective integrator.

References

[1] C. W. Gear, I. G. Kevrekidis, *Projective methods for stiff differential equations: Problems with gaps in their eigenvalue spectrum*, SIAM J. Sci. Comput. 24 (2003) pp. 1091–1106.

[2] B. P. Sommeijer, L. F. Shampine, J. G. Verwer, *RKC: An explicit solver for parabolic PDEs*, J. Comput. Appl. Math. 88 (1998), pp. 315–326.

[3] P. J. van der Houwen, *The development of Runge-Kutta methods for partial differential equations*, Appl. Num. Math. 20 (1996) pp. 261–272.

Degenerate hyperbolic systems in heat exchanger modelling: Analysis and numerical approximation

Michael Hanke

(joint work with Magnus Strömgren)

In a recent physical experiment with carbon dioxide as a refrigerant in a heat pump system two different steady states with very different coefficients of performance were observed. Numerical experiments to explain the behavior of the heat pump using standard system simulation tools failed very often. The heat exchanger may be modelled by the one-dimensional Euler equations of compressible fluid flow. [3] The physical properties of the flow and the refrigerant restrict the choice of the primary variables in the mathematical model. For one-component two-phase systems, the choice of pressure p and specific enthalpy h as thermodynamic variables is advantageous since they uniquely determine the state. The more common choice of pressure and temperature T is not suitable since the pressure is determined by the temperature alone in the two-phase region. As a third variable it is convenient to use the mass flow rate F. Since the flow in the heat pump system is characterized by very low Mach numbers everywhere exept in the expansion valve, we can eliminate the time scales associated with sound waves. This leads to a reduced model, the zero Mach-number limit,

$$\begin{aligned} \mathrm{A}\frac{\partial \rho}{\partial t} + \frac{\partial F}{\partial z} &= 0, \\ A\frac{\partial p}{\partial z} &= R, \\ A\frac{\partial e}{\partial t} + \frac{\partial}{\partial z}(Fh) &= Q. \end{aligned}$$

Here,the internal energy per unit volume is $e = \rho h - p$. The system is closed by the constitutive relations $\rho = f(p,h)$ and $T = g(p,h)$. The right-hand sides R and Q describe friction and the heat exchange, respectively. They are given by

$$R = -L_f \frac{F^2}{A\rho}\operatorname{sign} F \quad Q = A_{\mathrm{exch}}\alpha(T_{\mathrm{ext}} - T).$$

$A, A_{\mathrm{exch}}, T_{\mathrm{ext}}$ and L_f are given positive constants. In the friction-free case, $L_f = 0$.

The final system does no longer contain time derivatives in the momentum equations. Therefore, this degenerated hyperbolic system can be considered as

a partial differential-algebraic equation (PDAE). The following questions are of interest:

- Is the model well-posed?
- What are the correct boundary conditions?
- What is the index of this PDAE?
- How should one discretize the system?

We will answer these questions for the linearized frozen coefficient problem

$$\mathbf{A}\mathbf{u}_t + \mathbf{B}\mathbf{u}_x + \mathbf{C}\mathbf{u} = g(x,t)$$

where $\mathbf{u} = (F, p, h)^T$ and

$$\mathbf{A} = \begin{pmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & 0 \\ 0 & a_{32} & a_{33} \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{pmatrix}$$

and $\mathbf{C}$ is a general matrix. In the frictionless case ($L_f = 0$) it holds $c_{21} = c_{22} = c_{33} = 0$ while, for $L_f > 0$, we have $0 < \alpha_0 \leq c_{21}/L_f \leq \alpha_1$. In order to gain some insight into the structure of the system, the matrix pencil $(\mathbf{A}, \mathbf{B})$ is transformed into its Kronecker canonical form:

$$\begin{pmatrix} 1 & & \\ & 0 & 0 \\ & 1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}_t + \begin{pmatrix} u & & \\ & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}_z + \mathbf{D} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} \tilde{g}_1 \\ \tilde{g}_2 \\ \tilde{g}_3 \end{pmatrix}.$$

Here, u denotes the speed of the flow. Hence, we have one characteristic left from the Euler equations, as expected. If there wouldn't be any source term, then $\mathbf{D} = 0$. This case is discussed in detail in [4]. The time (differentiation) index becomes 2, while the space index is 0.

Assume now that friction as well as the other source terms are present. Then it can be shown that the system consists of

- a hyperbolic equation for v_1,
- a parabolic equation for v_2,
- and a differential relation for v_3.

Moreover, the following energy estimate can be proven,

$$\begin{aligned} \|\mathbf{v}(\cdot,t)\| \quad \leq \quad & C(t)\{\int_0^t (\|\tilde{g}(\cdot,\tau)\|^2 + \|\tilde{g}_x(\cdot,\tau)\|^2)d\tau \\ & + \|v_1(\cdot,0)\|^2 + \|v_2(\cdot,0)\|^2 + \|v_{1,x}(\cdot,0)\|^2 + \|v_{2,x}(\cdot,0)\|^2\} \end{aligned}$$

provided that the correct boundary conditions are given. They can be interpreted in terms of the original physical quantities:

- h and p at inflow and F at outflow (if $d_{33} < d_{22}$)
- h and F at inflow and p at outflow (if $d_{33} > d_{22}$)

The estimate can be interpreted in terms of a perturbation index: it is 1 with respect to time while it becomes 2 with respect to space! So the introduction of the friction term decreases the time index as intended but increases at the same time the space index.

The energy estimate above was derived under the assumption that Dirichlet boundary conditions are given for the *transformed* variables. In fact, the boundary conditions are posed in terms of the *physical* variables $\mathbf{u} = (F, p, h)^T$. Dirichlet boundary conditions for $\mathbf{u}$ lead to a coupling of the boundary conditions for $\mathbf{v}$. Does this have any influence on the stability properties? It can be shown that the linearized system is stable if the Reynolds number is not too large. It is not known if this restriction is really necessary or not for the system considered.

Having in mind that the system is a part of a larger network it is appropriate to discretize it by the method of lines (MOL). This discretization is done in the *physical* variables. By transforming it to the canonical ones the latter become coupled. For a toy problem, by using a simple upwind discretization the following can be shown:

If the step size is restricted by $\Delta x < Cu$, then

- the numerical scheme becomes weakly unstable;
- the resulting differential-algebraic equation has the tractability index 1.

Thus, for sufficiently small step sizes, the properties of the continuous system are resembled.

References

[1] M. Hanke, K.H.A. Olsson, Magnus Strïgren: *Stability analysis of a degenerate hyperbolic system modeling a heat exchanger.* Submitted 2005

[2] M. Hanke, M. Strïgren: *MOL discretization of a degenerate hyperbolic PDAE system.* Manuscript 2006

[3] P.G. Mehta, B. Eisenhower, J. Oppelstrup: *Bifurcation analysis of the CO_2 steady state model.* Internal Report, United Technologies Research Center, Hartford, CT, 2002

[4] W. S. Martinsson, P. I. Barton: Index and characteristic analysis of linear PDE systems. *SIAM J. Sci. Comput.,* 24(3):905-923, 2002

Exponential integrators of Rosenbrock-type

Marlis Hochbruck and Alexander Ostermann

(joint work with Julia Schweitzer)

1. Introduction

We consider a system of ordinary differential equations in autonomous form

$$y'(t) = f(y), \qquad y(t_0) = y_0, \tag{1}$$

assuming that the linearisation $J = Df(y)$ is uniformly sectorial in a neighbourhood of the exact solution. Consequently, there exist constants C and ω (both independent of y) such that

$$\left\| \mathrm{e}^{tJ} \right\| \le C\, \mathrm{e}^{\omega t}, \qquad t \ge 0. \tag{2}$$

Typical examples are abstract nonlinear parabolic equations, see [4], and their spatial discretisations.

Recently, we studied a class of explicit exponential Runge–Kutta methods for similar problems, see [2]. Due to the involved structure of the order conditions, however, it seems to be difficult to construct reliable and efficient error estimates for these methods. Moreover, in contrast to classical time integrators, exponential Runge–Kutta methods are *not invariant* under linearisation. This results in an error behaviour similar to classical W-methods, see [1]. Therefore, one has to expect large errors whenever the linear part is not well chosen.

2. Method class

Motivated by the observations just mentioned, we propose to linearise the right-hand side of (1) in each step, as it is done in classical Rosenbrock methods. Thus we write

$$y' = J_n y + g_n(y), \qquad J_n = Df(y_n), \quad g_n(y) = f(y) - J_n y. \tag{3}$$

Here, y_n is the numerical approximation to $y(t_n)$.

Applying then an exponential Runge–Kutta method to (3) gives the following s-stage exponential Rosenbrock-type scheme

$$Y_{ni} = \mathrm{e}^{c_i h J_n} y_n + h \sum_{j=1}^{i-1} a_{ij}(hJ_n)\, g_n(Y_{nj}), \tag{4a}$$

$$y_{n+1} = \mathrm{e}^{h J_n} y_n + h \sum_{i=1}^{s} b_i(hJ_n)\, g_n(Y_{ni}). \tag{4b}$$

For a variable step size implementation of (4), we base the step size selection on a local error control. For that purpose, we consider the embedded error estimator

$$\widehat{y}_{n+1} = \mathrm{e}^{h J_n} y_n + h \sum_{i=1}^{s} \widehat{b}_i(hJ_n)\, g_n(Y_{ni}) \tag{5}$$

and take the difference $\|y_{n+1} - \widehat{y}_{n+1}\|$ as error estimate.

3. Stiff order conditions

As usual in exponential integrators, the functions

$$\varphi_k(hJ) = h^{-k} \int_0^h \mathrm{e}^{(h-\tau)J} \frac{\tau^{k-1}}{(k-1)!}\, \mathrm{d}\tau, \qquad k \geq 1$$

play an important role. For sectorial operators J, the bound (2) shows that these functions are well defined and bounded on compact time intervals.

It is shown in [3] that the stiff order conditions for an exponential Rosenbrock-type method are given as:

No.	Order	Order Condition
1	1	$\sum_{i=1}^{s} b_i(hJ) = \varphi_1(hJ)$
2	2	$\sum_{j=1}^{i-1} a_{ij}(hJ) = c_i\varphi_1(c_i hJ), \quad 2 \le i \le s$
3	3	$\sum_{i=2}^{s} b_i(hJ)c_i^2 = 2\varphi_3(hJ)$
4	4	$\sum_{i=2}^{s} b_i(hJ)c_i^3 = 6\varphi_4(hJ)$

The first, third, and fourth order condition are just the (exponential) quadrature conditions, the second one is the well-known $C(1)$ condition, generalised to the operator case.

It is worth noting that the exponential Euler method applied to (3) is second-order accurate. It has one stage ($s = 1$) with weight $b_1(hJ) = \varphi_1(hJ)$ and consequently satisfies the first two order conditions.

4. Examples

From the above order conditions, it is straightforward to construct pairs of embedded methods up to order 4. We consider two examples. The method `exprb32` is a third-order method with a second-order error estimator (the exponential Euler method). Its coefficients are

$$\begin{array}{c|cc} c_1 & & \\ c_2 & a_{21} & \\ \hline & b_1 & b_2 \\ & \widehat{b}_1 & \end{array} \quad = \quad \begin{array}{c|cc} 0 & & \\ 1 & \varphi_1 & \\ \hline & \varphi_1 - 2\varphi_3 & 2\varphi_3 \\ & \varphi_1 & \end{array}$$

The method `exprb43` is a fourth-order method with a third-order error estimator. Its coefficients are

$$\begin{array}{c|ccc} c_1 & & & \\ c_2 & a_{21} & & \\ c_3 & a_{31} & a_{32} & \\ \hline & b_1 & b_2 & b_3 \\ & \widehat{b}_1 & \widehat{b}_2 & \widehat{b}_3 \end{array} \quad = \quad \begin{array}{c|ccc} 0 & & & \\ \frac{1}{2} & \frac{1}{2}\varphi_1\left(\frac{1}{2}\cdot\right) & & \\ 1 & 0 & \varphi_1 & \\ \hline & \varphi_1 - 14\varphi_3 + 36\varphi_4 & 16\varphi_3 - 48\varphi_4 & -2\varphi_3 + 12\varphi_4 \\ & \varphi_1 - 14\varphi_3 & 16\varphi_3 & -2\varphi_3 \end{array}$$

5. Stability and Convergence

For proving convergence estimates, the temporal smoothness of the *exact solution* is one of our basic ingredients. Using this property, we establish by Taylor series expansion a recursion for the global errors $E_n = y_n - y(t_n)$ in terms of the defects

$$E_{n+1} = \mathrm{e}^{h_n J_n} E_n + h_n \Delta_n.$$

Here Δ_n depends on the defects and on E_n itself. The stability of this recursion is all-important. It is verified in [3] that there exist constants C and Ω such that

$$\left\| \mathrm{e}^{h_n J_n} \cdots \mathrm{e}^{h_0 J_0} \right\| \;\leq\; C\, \mathrm{e}^{\Omega(h_0+\ldots+h_n)},$$

whenever the involved step sizes are sufficiently small. We emphasise that our proof of this result does *not* require the unrealistic condition $\left\| \mathrm{e}^{h_m J_m} \right\| \leq 1$.

A method is said to have *order* p, if it fulfils the stiff order conditions up to order p. For such methods, we have the following convergence result, see [3].

Theorem (Convergence). *Under the above assumptions, for $H > 0$ sufficiently small and $T \geq t_0$, there exists a constant C such that the global error satisfies*

$$\left\| y_n - y(t_0 + nh) \right\| \;\leq\; C\, h^p,$$

uniformly for all $0 < h \leq H$ and all $n \geq 0$ with $nh \leq T - t_0$. The constant C is independent of n and h.

Methods up to order 4 can be constructed easily, see the previous section. For numerical comparisons, we refer to [3].

References

[1] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*. 2nd rev. ed., Springer, Berlin, 1996.

[2] M. Hochbruck and A. Ostermann, *Explicit exponential Runge-Kutta methods for semilinear parabolic problems*. SIAM J. Numer. Anal. **43** (2005), 1069–1090.

[3] M. Hochbruck, A. Ostermann, and J. Schweitzer, *Exponential Rosenbrock-type methods*. In preparation.

[4] A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems*. Birkhäuser, Basel, 1995.

Optimization of DAEs with applications in Optimal Control

Shivakumar Kameswaran

(joint work with Lorenz T. Biegler)

Dynamic optimization aims at optimizing systems that are governed by differential equations. From a mathematical viewpoint, a dynamic optimization problem is an optimal control problem, which formally refers to the minimization of a cost (objective) function subject to constraints that represent the dynamics of the system. The last decade has witnessed a tremendous amount of effort going into optimization of DAEs. The focus was on developing numerical algorithms and optimization platforms, and solving interesting applications.

In order to cater to the scale and the complexity of present-day applications, the following directions must be explored: design of powerful numerical methods, optimization of systems governed by PDEs, ability to handle discrete decisions, identification of problem classes that can be solved by various dynamic optimization methodologies, reliability of NLP methodologies for dynamic optimization,

ill-conditioned systems and model reduction. My Ph.D. research aims at addressing some of these issues using rigorous theoretical tools and/or characteristic examples, and at the same time, use the results for solving large-scale industrial applications to realize the benefits.

In collaboration with Prof. Biegler, I have addressed the following issues:

- ***Discretize then Optimize* vs. *Optimize then Discretize*:** Research in this direction has focused on classification of problems based on whether it is advantageous to discretize all the dynamic constraints and then solve the large-scale NLP, or to discretize the optimality conditions of the original dynamic optimization problem. If the dynamic optimization problem is well-posed, then we have demonstrated the equivalence between the two approaches for a class of discretization schemes. Our analysis has also been instrumental in devising numerical procedures for dynamic optimization problems with high-index path constraints and singular optimal control problems. We have successfully applied our results for the boundary control of a heat transfer problem, and optimal control of fed-batch bioreactors and semi-continuous chemical reactors.
- **Reliability of NLP-Based Methods for Dynamic Optimization:** Although NLP-based methods have been used for an entire decade for the solution of dynamic optimization problems, characterizing (as a function of the step size) the relationship between the NLP solution and the solution of the original dynamic optimization problem is still an active area of research. In this direction, we have addressed convergence rates for NLP-based methods, and have identified classes of problems for which the reliability of NLP-based methods can be proved rigorously. Our results have applications in adjoint estimation, error analysis, and mesh refinement. We have also demonstrated the implication of our results on the temperature control of a batch reactor.
- **Handling Discrete Decisions:** A number of applications of dynamic optimization also possess discrete components. These discrete components introduce discontinuity into the optimization problem. Integer variables can be used to model these discontinuities, but the problem then becomes combinatorially expensive. We have overcome this difficulty by using complementarity conditions for modeling certain discrete decisions. The advantage of this formulation is that it does not use integer variables, and thus an NLP solver can be employed rather than a specialized mixed-integer nonlinear programming solver. We have used complementarity formulations for modeling and solving a number of interesting applications in optimal control, reservoir engineering, and chemical engineering.
- **Large-scale Parameter Estimation for a Reservoir Application:** In a project funded by ExxonMobil Upstream Research Company, we have developed a novel complementarity-based procedure for the estimation of

relative permeability and capillary pressure functions, from experimental data on oil-reservoir core samples. The values of these flow-functions are crucial for proper exploitation of petroleum resources. The coupled PDEs that are used to model two-phase flow phenomena, and boundary conditions that switch in a discrete manner depending on the value of the capillary pressure at the boundary make it a challenging optimization problem. Spatial and temporal discretization of the PDEs results in a large-scale NLP, and we have successfully solved instances that involve about 89000 variables and constraints. The solution procedure for this problem has benefited heavily from our research on the aforementioned directions.

- **Trajectory Planning for Fuel Cell/Gas Turbine Power Generation Systems:** In a different project funded by FuelCell Energy, Inc. we have developed a methodology for trajectory planning of hybrid power generation systems to achieve better control performance than conventional control. The goal is to predict controller moves to meet the power requirements, and this is crucial for the operation of such plants. We have formulated this as a dynamic optimization problem, and have successfully applied our research on dynamic optimization to solve the resulting large-scale problem (60000 variables and constraints).

We believe that the future of dynamic optimization lies in large-scale applications. Also, this decade is marked by increased efforts in modeling and simulation of biological and nano-scale processes/phenomena; optimization is a natural extension. With research efforts in the aforementioned directions, dynamic optimization is emerging as a much sought after tool for such applications. Preprints of my papers can be found from my homepage `http://dynopt.cheme.cmu.edu/skk/skk.html`.

Numerical solution of large-scale optimal control problems in robust optimum experimental design

Ekaterina Kostina

(joint work with Hans Georg Bock, Stefan Körkel, Johannes P. Schlöder)

Estimating model parameters from experimental data is crucial to reliably simulate dynamic processes. In practical applications, however, it often appears that the experiments performed to obtain necessary measurements are expensive, but nevertheless do not guarantee sufficient identifiability. The optimization of one or more dynamic experiments in order to maximize the accuracy of the results of a parameter estimation subject to cost and other technical inequality constraints leads to very complex non-standard optimal control problems.

The problem of optimum experimental design can be described as follows: our aim is

- to design N_{exp} optimal experiments by choosing experimental variables $q = (q_1, ..., q_{N_{exp}})$ (e.g. initial concentrations, properties of the experimental device etc), and experimental controls $u = (u_1, ..., u_{N_{exp}})$ (e.g. temperature profiles of cooling/heating, inflow profiles etc)
- that minimize a "size" of a confidence region of parameters which is described by a suitable function Φ of a covariance matrix of the underlying parameter estimation problem

$$\min_{q,u} \quad \Phi(C(x, p, q, u))$$

- such that the state variables $x = (x_1, ..., x_{N_{exp}})$, parameters p, design parameters $q = (q_1, ..., q_{N_{exp}})$ and design controls $u = (u_1, ..., u_{N_{exp}})$ satisfy control and state constraints

$$c_i(t, x_i(t), p, q_i, u_i(t)) \geq 0, \ t \in T_i = [t_0^i, t_f^i], \ i = 1, 2, ..., N_{exp},$$

- and $x_i(t) = (y_i(t), z_i(t)), q_i, u_i(t), p, \ t \in T_i$, satify model dynamics, e.g.

$$\begin{array}{rcll} \dot{y_i}(t) & = & f_i(t, y_i(t), z_i(t), p, q_i, u_i(t)), & \\ 0 & = & g_i(t, y_i(t), z_i(t), p), & t \in T_i, \end{array} \quad i = 1, 2, ..., N_{exp},$$

and constraints of the underlying parameter estimation problem

$$r_i(y_i, p, q_i, u_i) = 0, \ i = 1, 2, ..., N_{exp}.$$

One of the difficulties is that the objective function is a function of a covariance matrix and therefore already depends on a generalized inverse of the Jacobian of the underlying nonlinear parameter estimation problem, see [11, 6]. While experimental design for linear models is well established and discussed, e.g. in [1, 9, 13], numerical methods for design of experiments for nonlinear dynamic systems where first developed in [2, 3, 11, 12]. The numerical methods are based on the direct approach, according to which the control functions are parameterized on an appropriate grid by local support functions, the solution of the DAE systems and the state constraints are discretized. As a result, we obtain a finite-dimensional constrained nonlinear optimization problem which can be formally written as a general nonlinear programming problem

$$(1) \qquad \min_{\xi \in \mathbb{R}^{n_\xi}} \varphi(\xi, s, p) \ \text{ s.t. } \psi_i(\xi, s, p) = 0, i \in \mathcal{E}, \psi_i(\xi, s, p) \leq 0, i \in \mathcal{I}.$$

Here we summarize all experimental design variables, the (parameterized) controls in the n_ξ vector ξ, $s \in \mathbb{R}^{n_s}$ denotes a parameterization of the state variables of the process models, p is an n_p-vector of parameters, the functions $\varphi : \mathbb{R}^{n_\xi + n_p} \to \mathbb{R}$, $\psi_i : \mathbb{R}^{n_\xi + n_p} \to \mathbb{R}$, $i \in \mathcal{E} \cup \mathcal{I}$, are twice-continuously differentiable. Note, that the equality constraints $\psi_i(\xi, s, p) = 0, i \in \mathcal{E}$, contain the discretized boundary value problem. The problem (1) is solved by an SQP method.

The main effort for the solution of the optimization problem by the SQP method is spent on the calculation of the values of the objective function and the constraints as well as its gradients. Efficient methods for derivative computations combining internal numerical differentiation [4] of the solution of the DAE and automatic differentiation of the model functions [10] have been developed. For

more detailed discussion of the numerical methods for nonlinear optimum experimental design see [2, 3, 11, 12].

As we can see the experimental design optimization problem is formulated for the assumed parameter values which are, however, only known to lie in a possibly large confidence region. Our next aim is to construct robust experiments that is experiments that are less sensitive to parameter uncertainty.

We discuss here our approach for a general optimization problem (1). We assume that the parameters are only known to lie in a region defined by a "ball" around the nominal parameter values p_0

$$p \in \mathcal{P} := \{p : ||Q(p - p_0)|| \leq \gamma\},$$

where Q is a given nonsingular matrix. Our aim is to find a solution ξ^* which is robust, i.e. insensitive, to "small" perturbations in p. For this purpose, following one of the classical approaches since middle of the sixties, we may form and solve the worst-case design problem:

$$\begin{array}{lll} \min\limits_{\xi} \max\limits_{p\in\mathcal{P}} & \varphi(\xi, s, p), \\ \text{s.t.} & \psi_i(\xi, s, p) = 0, i \in \mathcal{E}, \forall p \in \mathcal{P} \\ & \max\limits_{p\in\mathcal{P}} \psi_i(\xi, s, p) \leq 0, i \in \mathcal{I}. \end{array} \tag{2}$$

The optimization problem (2) is a semi-infinite programming problem. The solution methods for such problems require the determination of global optima of nonlinear subproblems which may be computationally extremely expensive. In order to compute robust solutions we suggest to approximate the worst-case problem (2) in the following way. First, we assume that we may reduce the problem to an inequality constrained one. For the sake of notation simplicity we assume that the number of constraints is equal to the number of the state variables s, the matrix

$$\frac{\partial \psi_{\mathcal{E}}(\xi, s_0, p_0)}{\partial s}, \psi_{\mathcal{E}}(\cdot) = \left(\begin{array}{c} \psi_i(\cdot) \\ i \in \mathcal{E} \end{array} \right)$$

is nonsingular at the pair p_0, s_0, satisfying $\psi_{\mathcal{E}}(\xi, p_0, s_0) = 0$, and there exist a sufficiently smooth function $s = s(p)$, $s(p_0) := s_0$, $p \in \mathcal{P}$, such that $\psi_{\mathcal{E}}(\xi, s(p), p) \equiv 0$, $p \in \mathcal{P}$. Denote

$$\mathcal{R} := \frac{\partial \psi_{\mathcal{E}}(\xi, s_0, p_0)}{\partial s}^{-1} \frac{\partial \psi_{\mathcal{E}}(\xi, p_0, s_0)}{\partial p}.$$

Then we approximate the problem (2) by

$$\begin{array}{lll} \min\limits_{\xi} \max\limits_{p\in\mathcal{P}} & \tilde{\varphi}(\xi, s_0, p_0) := \varphi(\xi, s_0, p_0) + \\ & \left(-\dfrac{\partial \varphi(\xi, s_0, p_0)}{\partial s} \mathcal{R} + \dfrac{\partial \varphi(\xi, s_0, p_0)}{\partial p} \right) (p - p_0), \\ \text{s.t.} & \max\limits_{p\in\mathcal{P}} \tilde{\psi}_i(\xi, s_0, p_0) := \psi_i(\xi, s_0, p_0) + \\ & \left(-\dfrac{\partial \psi_i(\xi, s_0, p_0)}{\partial s} \mathcal{R} + \dfrac{\partial \psi_i(\xi, s_0, p_0)}{\partial p} \right) (p - p_0) \leq 0, \ i \in \mathcal{I}, \end{array} \tag{3}$$

using Taylor expansions for the functions φ, ψ_i, $i \in \mathcal{I}$, with respect to p. The inner problems in (3) are the maximizations of linear functions subject to convex constraints which can be solved explicitly. Using the explicit solution of the inner problems in (3), we may rewrite the approximate worst-case problem (3) as follows

$$
\begin{aligned}
(4)\ \min_{\xi} \quad & \varphi(\xi, s_0, p_0) + \gamma ||Q^{-T} \left(-\mathcal{R}^T \frac{\partial \varphi(\xi, s_0, p_0)^T}{\partial s} + \frac{\partial \varphi(\xi, s_0, p_0)^T}{\partial p} \right) ||_*, \\
\text{s.t.} \quad & \psi_i(\xi, p_0) + \gamma ||Q^{-T} \left(-\mathcal{R}^T \frac{\partial \psi_i(\xi, s_0, p_0)^T}{\partial s} + \frac{\partial \psi_i(\xi, s_0, p_0)^T}{\partial p} \right) ||_* \leq 0, \\
& i \in \mathcal{I},
\end{aligned}
$$

where $||\cdot||_*$ denotes a dual norm to $||\cdot||_\nu$. The second term in the cost function and the constraints can be interpreted as a penalty for uncertainty in the parameters.

Applying SQP-type method for solving (4), we need second-order derivative of functions ψ_i. However, in case of Euclidian norm we may compute the necessary derivatives very efficiently. Indeed, in this case we need directional derivatives of the form $\dfrac{\partial^2 \psi_i(\xi, s_0, p_0)}{\partial p \partial \xi} \dfrac{\partial \psi_i(\xi, s_0, p_0)^T}{\partial p}$ which can be computed by means of automatic differentiation. For computing the sensitivities $\mathcal{R}$ one may again apply methods of automatic differentiation.

The methods of robust optimal experimental design have been applied to several applications in chemistry [7] and bio-chemistry [5] and they have allowed to estimate reliably unknown parameters and to reduce significantly experimental costs. The methods for robust nonlinear optimization presented were applied also for solving optimal control problems, see [8].

References

[1] A. C. Atkinson, A. N. Donev, *Optimum Experimental Designs*, Oxford University Press (1992)

[2] I. Bauer, H. G. Bock, S. Körkel, J. P. Schlöder, *Numerical methods for initial value problems and derivative generation for DAE models with application to optimum experimental design of chemical processes*, in: Keil, F., Mackens, W., Voss, H., Werther, J. (eds) Scientific Computing in Chemical Engineering II, **2**, Springer, Berlin-Heidelberg (1999), 282 289.

[3] I. Bauer, H. G. Bock, S. Körkel, J. P. Schlöder, *Numerical methods for optimum experimental design in DAE systems*, Journal of Computational and Applied Mathematics, **120** (2000), 1–25

[4] H. G. Bock, *Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen*, Bonner Mathematische Schriften, **187**, Bonn (1987)

[5] H. G. Bock, S. Körkel, E. A. Kostina, J. P. Schlöder, *Methods for design of optimal experiments with application to parameter estimation in enzyme catalytic processes*, in: Hicks, M. G., Kettner C. (eds) Experimental Standard Conditions of Enzyme Characterizations, Proccedings of the International Beilstein Workshop, Beilstein-Institut zur Förderung der Chemischen Wissenschaften, 45 –70 (2004)

[6] H. G. Bock, E. A. Kostina, O. I. Kostyukova, *Covariance matrices for constrained parameter estimation problems*, accepted in SIAM Journal on Matrix Analysis and Applications (2006)

[7] H. G. Bock, S. Körkel, E. A. Kostina, J. P. Schlöder, *Numerical methods for optimal control problems in design of robust optimal experiments for nonlinear dynamic processes*, Optimization Methods and Software, **19**, issue 3-4, (220), 327 – 338.

[8] M. Diehl, H. G. Bock, E. A. Kostina, *An approximation technique for robust nonlinear optimization*, Mathematical Programming, **107** (2006), 213-230.
[9] V. V. Fedorov, *Theory of Optimal Experiments*, Probability And Mathematical Statistics, Academic Press, London (1972)
[10] A. Griewank, *Evaluating Derivatives. Principles and Techniques of Algorithmic Differentiation*, Frontiers in Applied Mathematics, SIAM (2000)
[11] S. Körkel, *Numerische Methoden für Optimale Versuchsplanungsprobleme bei nichtlinearen DAE-Modellen*, PhD thesis, Universität Heidelberg (2002)
[12] S. Körkel, E. A. Kostina, *Numerical methods for nonlinear experimental design,* in: Bock, H. G., Kostina E. A., Phu H. X., Rannacher R. (eds) Modeling, Simulation and Optimization of Complex Processes, Proceedings of the International Conference on High Performance Scientific Computing, 2003, Hanoi, Vietnam, Springer (2004)
[13] F. Pukelsheim, *Optimal Design of Experiments*, John Wiley & Sons, Inc., New York (1993)

Some problems connected with linear-quadratic optimal control problems for descriptor systems

Galina Kurina
(joint work with Roswitha März)

I. Index criteria for differential algebraic equations arising from linear-quadratic optimal control problems [1]**.** We consider the quadratic cost functional

$$J(u,x)=\frac{1}{2}\langle x(T),Vx(T)\rangle+\frac{1}{2}\int_{t_0}^{T}\left\langle\begin{pmatrix}x\\u\end{pmatrix},\begin{pmatrix}W & S\\ S^* & R\end{pmatrix}\begin{pmatrix}x\\u\end{pmatrix}\right\rangle dt$$

to be minimized on solutions of the linear differential algebraic equation (DAE) with properly stated leading term

$$A(Bx)'=Cx+Du,\quad x=x(t)\in\mathbb{R}^m,\quad u=u(t)\in\mathbb{R}^l, \tag{1}$$

subject to the initial condition

$$A(t_0)B(t_0)x(t_0)=z_0.$$

The coefficients are supposed to be continuous matrix functions with $A(t)\in L(\mathbb{R}^n,\mathbb{R}^k), B(t)\in L(\mathbb{R}^m,\mathbb{R}^n), C(t)\in L(\mathbb{R}^m,\mathbb{R}^k), D(t)\in L(\mathbb{R}^l,\mathbb{R}^k), W(t)\in L(\mathbb{R}^m),\ S(t)\in L(\mathbb{R}^l,\mathbb{R}^m), R(t)\in L(\mathbb{R}^l);\ V\in L(\mathbb{R}^m);\ V^*=V, W^*(t)=W(t), R^*(t)=R(t); V$ and $\begin{pmatrix}W(t) & S(t)\\ S(t)^* & R(t)\end{pmatrix}$ are positive semidefinite, $t\in[t_0,T]$.

The system

$$\begin{array}{rcl}A(Bx)' & = & Cx+Du,\\ -B^*(A^*\psi)' & = & Wx+C^*\psi+Su,\\ 0 & = & S^*x+D^*\psi+Ru,\end{array} \tag{2}$$

following from the control optimality condition, is regular with tractability index 1 if and only if the matrices

$$\begin{pmatrix}AB-CQ_0 & D\end{pmatrix},\quad\begin{pmatrix}G_0^*-C^*Q_{*0} & WQ_0 & S\\ -D^*Q_{*0} & S^*Q_0 & R\end{pmatrix}$$

have full row rank on the interval $[t_0, T]$. Here $G_0 = AB$; Q_0, Q_{*0} are the orthogonal projectors onto $KerG_0, KerG_0^*$ respectively.

It should be noted that the DAE (1) to be controlled is not assumed to be regular or to have an index.

The definition of self-adjoint DAEs is introduced. The conditions ensuring the Hamiltonian structure of inherent ODEs are established for self-adjoint DAEs.

For the controlled DAE consider a properly stated leading term with $\ker A(t) = \{0\}$. For the system (2) suppose the conditions assuring the regularity with index 1 to be fulfilled. Then the inherent ODE of (2) has the Hamiltonian structure

$$z' = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} Ez, \quad z := \begin{pmatrix} Bx \\ -A^*\psi \end{pmatrix}, \quad E(t)^* = E(t).$$

This property is very useful for the solvability of boundary value problems. The example from [1] shows the significance of the condition concerning the leading term.

Let $m = k$, DAE (1) be regular with index zero or with index one, and $imD \subseteq imAB$, $imS \subseteq im(AB)^*$, then the DAE (2) is regular with tractability index one if R is invertible and vice versa.

If R is singular the DAE (2) can not be regular with index two in general case. Reasonable conditions yield the regularity with tractability index three for (2).

II. Feedback approach using an implicit Riccati equation [2]**.** Let us assume that $R(t)$ is non-singular for $t \in [t_0, T]$ and consider the final value problem

$$B^*(A^*YB^-)'B = -Y^*C - C^*Y + (S + Y^*D)R^{-1}(S^* + D^*Y) - W, \qquad (3)$$

$$A(T)^*Y(T)B(T)^- = B(T)^{-*}VB(T)^-. \qquad (4)$$

If Y solves (3),(4) and $A^*YQ_0 = 0$ then $B^*A^*Y = Y^*AB$.

Let Y be a solution for (3),(4), the condition $A^*YQ_0 = 0$ be fulfilled and x_* be a solution of the IVP

$$A(Bx)' = Cx - DR^{-1}(S^* + D^*Y)x, \quad A(t_0)B(t_0)x(t_0) = z_0.$$

Then we get an optimal control in the feedback form $u_* = -R^{-1}(S^* + D^*Y)x_*$ and $J(u_*, x_*) = \frac{1}{2}\langle z_0, A(t_0)^{*-}B(t_0)^{-*}Y(t_0)^*z_0\rangle$.

The solution of the implicit Riccati equation is reduced to the solution of the algebraic Riccati equation and the standard differential Riccati equation resolved with respect to the derivative.

The solvability of the closed loop problem and the connection between the implicit Riccati equations and the implicit Hamiltonian systems have been also studied.

The another implicit Riccati equation was used by other authors earlier (see, for example, [3]). This another implicit Riccati equation fails for the different combinations of the coefficients in the example from [3], but the solution for (3) leads to an optimal control in a feedback form for this example.

III. Some non-standard linear-quadratic problems for descriptor systems [4]**.** We consider now the non-standard quadratic cost functional

$$J(u,y)=\frac{1}{2}\sum_{j=0}^{N+1}\langle y(t_j)-y_j,F_j(y(t_j)-y_j)\rangle+\frac{1}{2}\int_0^T\langle u(t),Ru(t)\rangle dt$$

to be minimized on the trajectories of a descriptor system

$$(Bx)'=Cx+Du,\quad y=Gx.$$

Here $t_0=0$, $t_{N+1}=T$, $0<t_1<\cdots<t_N<T$; t_j are fixed; $F_j=F_j^*\geq 0$, $R=R^*>0$, y_j are given. The coefficients are constant in this problem.

Two cases are researched, namely, when the initial value for the part of the state variable is given and when additional constraints for boundary points of the state variable are absent. For the third problem, output variable values in the fixed points are given. The solvability of these optimal control problems has been proved.

Adjoint variables and optimal controls are discontinuous functions when $t=t_j$ in general case. The jumps for the adjoint variable $\psi(\cdot)$ are given by the formulas

$$B^*(\psi(t_j-0)-\psi(t_j+0))=-G^*F_j(y_*(t_j)-y_j),$$

$y_*(t)=Gx_*(t)$, $x_*(\cdot)$ is a trajectory corresponding to the control

$$u_*(t)=R^{-1}D^*\psi(t),B^*(\psi(t))'=-C^*\psi(t),\quad t\neq t_j.$$

IV. Discrete problems [5]**,** [6]. Let us consider the problem of minimizing the quadratic functional

$$J(u,x)=\frac{1}{2}\langle x_N,Vx_N\rangle+\frac{1}{2}\sum_{i=0}^{N-1}\left\langle\begin{pmatrix}x_i\\u_i\end{pmatrix},\begin{pmatrix}W_i&S_i\\S_i^*&R_i\end{pmatrix}\begin{pmatrix}x_i\\u_i\end{pmatrix}\right\rangle$$

on trajectories of a descriptor system

$$A_{i+1}B_{i+1}x_{i+1}=C_ix_i+D_iu_i,\quad i=\overline{0,N-1},\quad A_0B_0x_0=z_0.$$

Under some conditions, the implicit system, following from the control optimality condition, provides an explicit nonnegative standard Hamiltonian system for the pair $(B_ix_i,A_i^*\psi_i)$ and the considered optimal control problem is solvable.

We put $B_i\equiv I$. If the symmetric operators K_i are the solution of the problem

$$A_i'K_iA_i=W_i+C_i'K_{i+1}C_i-(S_i+C_i'K_{i+1}D_i)L_i(S_i+C_i'K_{i+1}D_i)',$$
$$A_N'K_NA_N=V$$

such that the operators

$$L_i^{-1}=R_i+D_i'K_{i+1}D_i$$

are positive definite, then $u_i^*=-L_i(S_i'+D_i'K_{i+1}C_i)x_i$ is an optimal control in a feedback form.

In general case, the implicit discrete operator Riccati equation has no a symmetric non-negative definite solution but it has a solution ensuring the positive definiteness of the operators L_i.

This work was partially supported by the RFBR (project No.06-01-00296-a).

References

[1] K. Balla, G.A. Kurina, R. März, *Index criteria for differential algebraic equations arising from linear-quadratic optimal control problems*, Institut für Mathematik an der Mathematisch-Naturwissenschaftlichen Fakultät II der Humboldt-Universität zu Berlin (Preprint; 2003-**14**) (2003).

[2] G.A. Kurina, R. März, *Feedback solutions of optimal control problems with DAE constraints*, Institut für Mathematik an der Mathematish-Naturwissenschaftlichen Fakultät II der Humboldt-Universität zu Berlin (Preprint; 2005-**9**)(2005).

[3] P. Kunkel, V. Mehrmann, *The linear quadratic optimal control problem for linear descriptor systems with variable coefficients*, Math. Control Signals Systems **10** (1997), 247–264.

[4] G. Kurina, *On some linear-quadratic optimal control problems for descriptor systems*, Research Reports in Mathematics, Department of Mathematics, Stockholm University, no.1 (2006).

[5] G. Kurina, *Linear-quadratic discrete optimal control problems for descriptor systems in Hilbert space*, Journal of Dynamical and Control Systems, **10**, no.3 (2004), 365–375.

[6] G.A. Kurina, *Feedback control for discrete descriptor systems*, Systems Science, **28**, no.2 (2002), 29–40.

Different Index Concepts, their Canonical Forms and Solvability of Linear DAEs

René Lamour

(joint work with Roswitha März)

The index of a (linear) DAE describes the number of differentiations of the right-hand side that are necessary to compute a solution. The generalization of the classical Weierstraß-Kronecker index for constant coefficient DAEs using the local matrix pencil leads to irrelevant results and therefore various index definitions have been provided.

We consider DAEs in standard form

$$E(t)x'(t) + F(t)x(t) = q(t), \tag{1}$$

or with properly stated leading term

$$A(t)(D(t)x(t))' + B(t)x(t) = q(t), \tag{2}$$

with sufficiently smooth (at least continuous) matrices $A(t) \in \mathbb{R}^{m\times n}$, $D(t) \in \mathbb{R}^{n\times m}$, $B(t), E(t), F(t) \in \mathbb{R}^{m\times m}$, $t \in I$.

It would be desirable to have a common canonical form for all concepts to compare the different definitions. However, up to now also the canonical forms are (in detail) different.

Our special interest (as a part in a project of the DFG Research Center MATHEON) was directed at the relation between the tractability index and the strangeness index.

Notice that the strangeness index is defined for equations (1). The tractabilty index is given for both formulations (1) and (2), see e.g. [4] and [6]. A projector

based version of the strangeness concept will be given in [3].
The standard canonical form (SCF) (see [1]) (also known as **S**teve **C**ampbell **F**orm)

$$\begin{aligned} u' + \mathcal{W}u &= L_u q, \\ \mathcal{N}w' + w &= L_w q, \end{aligned} \tag{3}$$

was derived for the tractability index [5] and for the strangeness index [2]. In contrast to the SCF definition of Campbell, who considered a strictly upper triangular matrix $\mathcal{N}$ with possible variable rank, constant ranks related to characteristic numbers of the tractability index or others of the strangeness index play an important role in the definition of both concepts. Here a rank change is considered a critical point.
The conjecture was that the characteristic numbers of both concepts are strongly related. Considering (3) for the tractability index and strangeness index the inner structure of the related nilpotent matrix

$$\mathcal{N} = \begin{pmatrix} 0 & \mathcal{N}_{01} & \cdots & \cdots & \mathcal{N}_{0,\mu-1} \\ & 0 & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & \mathcal{N}_{\mu-2,\mu-1} \\ & & & & 0 \end{pmatrix}$$

are different.
In the tractability index case the supdiagonal blocks have full column rank structure $\boxed{\mathcal{N}_{k,k+1}}$, but in the strangeness index case full row rank structure $\boxed{\mathcal{N}_{k,k+1}}$ with rank $s_{\mu-2-k}$.
Both index definitions are based on different matrix chains and therefore the idea of an equivalence proof concerning the rank conditions is to take the normal form of the first index and apply the matrix chain of the other one.

$$\begin{array}{ccc} \text{DAE of index } \mu & \xrightarrow{\text{Strangeness concept}} & \text{Normal form} \\ \text{Tractability}\big\downarrow\ \text{concept} & & \text{Tractability}\big\downarrow\ \text{concept} \\ \text{Normal form} & \xrightarrow{\text{Strangeness concept}} & \mu_t = \mu_s + 1? \end{array}$$

A regular DAE with strangeness index μ was proved to have tractability index μ+1.

Theorem 1: For (1) let the strangeness index be well defined. Let $(a_i, \bar{r}_i, s_i, u_i)$, $_{i=0,\ldots,\mu-1}$, denote the corresponding characteristic values, and let $s_{\mu-1} = 0,\ s_{\mu-2} \neq 0,\ u_i = 0,\ _{i=0,\ldots,\mu-1}$ (i.e., (1) has regular s-index $\mu - 1$).

Then, the DAE (1) is (in each proper reformulation) regular with tractability index μ and corresponding characteristic values $r_0 = \mathrm{rank} E = \bar{r}_0, \quad r_j = m - s_{j-1}$, $_{j=1,\ldots,\mu}$.

Up to the Oberwolfach conference we only believed the other direction to hold but thanks to the inspiring atmosphere of Oberwolfach we succeeded in proving the other direction.

Theorem 2: Let the regular strangeness index μ_s (see [3]) and the tractability index $\mu_t \geq 1$ be well defined for a linear DAE (2), then $\mu_t = \mu_s + 1$ and the characteristic values are given in a one-to-one relation.

Remarks: -Regular strangeness index means, that the DAE is not over or underdetermind or, using the characteristic values of the strangeness index concept that $u_i = 0$, $\forall i$.
-For the special case that $\mu_t = 0$ (i.e., we have an implicit regular ODE) the relation $\mu_t = \mu_s$ holds because of the different counting of the strangeness index.

References

[1] S. Campbell, *A General Form for Solvable Linear Time Varying Singular Systems of Differential Equations*, SIAM J. MATH. ANAL, 18 (1987) 1101-1115

[2] P. Kunkel and V. Mehrmann. *Canonical forms for linear differential-algebraic equations with variable coefficients.* J. Comput. Appl. Math., **56**(1994)225–259.

[3] R. Lamour, *A Projector Chain Representation of the Strangeness Index Concept*, in preparation

[4] R. März, *Numerical methods for differential-algebraic equations.* Acta Numerica, pp. 141-198, 1992.

[5] R. März, *Fine decouplings of regular differential algebraic equations*, Results in Mathematics, **46**(2004)57-72.

[6] R. März, *Solvability of linear differential algebraic equations with properly stated leading term*, Results in Mathematics, **45**(2004)88-105.

Stability radii for linear time-varying differential algebraic equations and their dependence on data

Vu Hoang Linh

(joint work with Nguyen Huu Du)

This research is concerned with the robust stability for time-varying systems of differential-algebraic equations (DAE-s) of the form

$$E(t)x'(t) = A(t)x(t), \quad t \geq 0, \tag{1}$$

where $E(\cdot) \in L_\infty^{loc}(0, \infty; \mathbb{K}^{n\times n})$ with absolutely continuous kernel, $\mathbb{K} = \{\mathbb{C}, \mathbb{R}\}$, and $A(\cdot) \in L_\infty^{loc}(0, \infty; \mathbb{K}^{n\times n})$. The leading term $E(t)$ is supposed to be singular for all

$t \geq 0$. We suppose that (1.1) generates an exponentially stable evolution operator $\Phi = \{\Phi(t,s)\}_{t,s\geq 0}$, i.e., there exist positive constants M and α such that

$$\|\Phi(t,s)\|_{\mathbb{K}^{n\times n}} \leq Me^{-\alpha(t-s)}, \quad t \geq s \geq 0. \tag{2}$$

We consider the system (1) subjected to structured perturbation of the form

$$E(t)x'(t) = A(t)x(t) + B(t)\Delta(C(\cdot)x(\cdot))(t), \quad t \geq 0 \tag{3}$$

where $B(\cdot) \in L_\infty(0,\infty;\mathbb{K}^{n\times m})$ and $C(\cdot) \in L_\infty(0,\infty;\mathbb{K}^{q\times n})$ are given matrices defining the structure of the perturbation and $\Delta : L_p(0,\infty;\mathbb{K}^m) \to L_p(0,\infty;\mathbb{K}^q)$ is an unknown disturbance operator which is supposed to be linear, dynamic, and causal. Thus, the system (3) represents a large class of linear functional differential equations including, e.g., delay equations, integro-differential equations, etc. In applications, the nominal system (1) plays the role of a simplified model problem, while the perturbed system (3) can be considered as a real-life problem.

The so-called stability radius is defined by the largest bound r such that the stability is preserved for all perturbations Δ of norm strictly less than r. Depending on $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, we talk about the complex stability radius or the real one. This measure of the robust stability was introduced by Hinrichsen and Pritchard [7]. Formulae of the stability radii for linear time-invariant systems of ordinary differential equations (ODE-s) with respect to time- and output-invariant, i.e., static perturbations can be found in [7, 11]. In lots of practical problems, uncertain perturbations may depend on the output feedback, as well. In [9], robust stability with respect to dynamic perturbations was considered for explicit time-invariant systems and a formula of the stability radius was given in term of the norm of a certain input-output operator. For time-varying systems, the most successful attempt for finding a formula of the stability radius was an elegant result given by Jacob [6]. On the other hand, systems occurring in various applications, such as optimal control, circuit design, multibody mechanics simulation, etc., are described by differential-algebraic systems, see [1]. Therefore, it is natural to extend the notion of the stability radius to differential-algebraic equations. The stability radius was formulated for linear time-invariant DAE-s of index-1, see [10, 2], and analyzed for implicit systems containing small parameters, see [3]. It is worth remarking that, as in the qualitative theory and in numerical methods for DAE-s, the index notion plays a very important role in the robust stability analysis, too.

The first aim of this research is to extend Jacob's result to time-varying systems (1) of index-1. We follow the tractability index approach proposed by Griepentrog and März [4]. First, a definition of the structured stability radii for (1) subjected to (3), denoted by $r_{\mathbb{K}}(E,A;B,C)$, is given. It is slightly different from the case of ODE-s that not only the stability, but also the index-1 property are required to be preserved. We propose the exact formula for $r_{\mathbb{K}}(E,A;B,C)$ as follows. Let $Q(\cdot)$ be an absolutely continuous projector onto $\ker E(\cdot)$. Set $P = I - Q$, $\overline{A} = A + EP'$ and $G = E - \overline{A}Q$. Assume that the following assumptions hold.

Assumption A1. System (1) is index-1 and there exist $\overline{M} > 0$, $\alpha > 0$ such that

$$\|\Phi_0(t,s)P(s)\| \leq \overline{M}e^{-\alpha(t-s)}, \quad t \geq s \geq 0.$$

Here $\Phi_0(t,s)$ is the Cauchy operator of the so-called inherent ODE system for (1).
Assumption A2. PG^{-1}, QG^{-1} and $Q_s := -QG^{-1}\overline{A}$ are essentially bounded on $[0,\infty)$.
We introduce the following operators

$$\begin{aligned}(\mathbb{L}_{t_0}u)(t) &= C(t)\int_{t_0}^{t}\Phi(t,s)PG^{-1}B(s)u(s)ds + CQG^{-1}B(t)u(t),\\ (\widetilde{\mathbb{L}}_{t_0}u)(t) &= CQG^{-1}B(t)u(t),\end{aligned}$$

for all $t \geq t_0 \geq 0$, $u \in L_p(0,\infty;\mathbb{K}^n)$, $p \geq 1$. The first operator is called the input-output operator associated with (1-3).

Theorem 1. *Let Assumptions A1-2 hold. Then*

$$r_{\mathbb{K}}(E,A;B,C) = \min\{\sup_{t_0\geq 0}\|\mathbb{L}_{t_0}\|^{-1}, \|\widetilde{\mathbb{L}}_0\|^{-1}\}.$$

If E is nonsingular, then by setting $Q=0$, one obtains Jacob's result. Furthermore, if E, A, B, and C are real, then the complex stability radius and the real one coincide. We note that the perturbed system may loose the index-1 property under the effect of perturbations. This yields an essential difference between the robust stability of a singular system and that of a regular one.

For time-invariant case, the following theorem extends the result for ODE-s by Hinrichsen et.al. [9] to DAE-s.

Theorem 2. *Suppose that the nominal time-invariant system (1) has index-1 and is exponentially stable. Then*

$$r_{\mathbb{K}}(E,A;B,C) = \|\mathbb{L}_0\|^{-1}.$$

Furthermore, if $p=2$, then

$$r_{\mathbb{C}}(E,A;B,C) = \left(\sup_{w\in i\mathbb{R}}\left\|C(wE-A)^{-1}B\right\|\right)^{-1}.$$

Further analysis is addressed to the dependence of the stability radii on data. This problem was investigated for time-invariant ODE-s [8], for time-varying ODE-s [5], and for time-invariant DAE-s [3]. Let $\{F_k(\cdot)\}_{k\in\mathbb{N}}$ be a sequence of measurable and essentially bounded matrix functions. We consider a sequence of the perturbed systems

$$E(t)x'(t) = (A(t)+F_k(t))\,x(t), \quad t\geq 0, \quad k=1,2,... \tag{4}$$

Assumption A3. With the projector function Q defined above, the matrices $\widetilde{G}_k = E-(\overline{A}+F_k)Q$ are invertible almost everywhere and for all k. Furthermore, $P\widetilde{G}_k^{-1}$, $Q\widetilde{G}_k^{-1}$ and $Q_{s,k} = -Q\widetilde{G}_k^{-1}\overline{A}$ are essentially bounded on $[0,\infty)$.

Theorem 3. *Let Assumptions A1-3 hold. Furthermore, suppose that the following assumptions hold:*

$$\begin{aligned}&i) \quad PG^{-1}F_k(\cdot)\in L_1(0,\infty;\mathbb{K}^{n\times n}), \quad \forall k\\ &ii) \quad \lim_{k\to\infty}\operatorname{ess\,sup}_{t\geq 0}\|P(\widetilde{G}_k^{-1}-G^{-1})(t)\| = 0,\\ &iii) \quad \lim_{k\to\infty}\operatorname{ess\,sup}_{t\geq 0}\|Q(\widetilde{G}_k^{-1}-G^{-1})(t)\| = 0.\end{aligned}$$

Then, the perturbed systems (4) remain index-1 and generate exponentially stable Cauchy operators, too. Moreover, their stability radii tend to those of the original (1), respectively, as k tends to infinity:

$$\lim_{k\to\infty} r_{\mathbb{K}}(E, A + F_k; B, C) = r_{\mathbb{K}}(E, A; B, C).$$

Further results and conclusions on data dependence are also obtained for the stability radii of special problems such as almost time-invariant systems and explicit time-varying systems. In addition, the dependence of the stability radii on the perturbation structure is analyzed, too. The publication of the results presented here, supplied with detailed proofs, is being in progress.

Acknowledgement. The speaker dedicates this talk to the memory of Professor Katalin Balla (1947–2005).

References

[1] K.E. Brenan, S.L. Campbell, L.R. Petzold, *Numerical solution of initial value problems in differential algebraic equations*, SIAM, Philadelphia, 1996.
[2] R. Byers, N.K. Nichols, *On the stability radius of generalized state-space systems*, Linear Algebra Applications, **188-189**(1993), 113–134.
[3] N.H. Du, V.H. Linh, *Robust stability of implicit linear systems containing a small parameter in the leading term*, IMA J. Mathematical Control Information, **23**(2006), 67–84.
[4] E. Griepentrog, R. März, *Differential-algebraic equations and their numerical treatment*, Teubner-Texte zur Mathematik, Leibzig, 1986.
[5] A. Ilchmann, I.M.Y. Mareels, *Stability radii for slowly time-varying systems*, in: Advances in mathematical system theory, Boston, Birkhäuser, 2001, 55–75.
[6] B. Jacob, *A formula for the stability radius of time-varying systems*, J. Differential Equations, **142**(1998), 167–187.
[7] D. Hinrichsen,A.J. Pritchard, *Stability for structured perturbations and the algebraic Riccati equation*, Systems Control Letters, **8**(1986), 105–113.
[8] D. Hinrichsen, A.J. Pritchard, *A note on some differences between complex and real stability radii*, Systems Control Letters, **14**(1990), 401–408.
[9] D. Hinrichsen, A.J. Pritchard, *Destabilization by output feedback*, Differential Integral Equations, **5**(1992), 2, 357–386.
[10] L. Qiu, E.J. Davison, *The stability robustness of generalized eigenvalues*, IEEE Transactions on Automatic Control, **37**(1992), 886–891.
[11] L. Qiu, B. Benhardson, A. Rantzer, E.J. Davison, P.M. Young, and J.C. Doyle, *A formula for computation of the real stability radius*, Automatica, **31**(1995), 879–890.

Symplectic integrators for general relativity

Christian Lubich

The Einstein equations of general relativity have a Hamiltonian formulation in a 3+1 slicing, as has been known since the work by Dirac and by Arnowitt, Deser and Misner nearly 50 years ago. The objective of this talk was to up possible ways to exploit the Hamiltonian structure in the numerical integration.

The Einstein equations have weak invariants, which can be viewed as constraints with zero Lagrange multipliers: the Hamiltonian and momentum constraints. They are, however, no longer automatically preserved under discretization. It turns out that by considering the shift as additional momentum variables and imposing a gauge condition, we obtain a formulation where the momentum constraints arise as hidden constraints of a holonomically constrained Hamiltonian system. The Rattle integrator is a very suitable integration method for this system: it is symplectic and enforces the momentum constraints, but it ignores the Hamiltonian constraint. Nevertheless, a result by Anderson and York implies that in the continuous equations there is no drift in the Hamiltonian constraint once the momentum constraints are enforced. Provided that this property carries over to the spatial semi-discretization, the Rattle integrator applied to the proposed formulation thus yields a very promising structure-preserving integration method.

References

[1] E. Hairer, C. Lubich, G. Wanner, *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer, Berlin, 2nd edition, 2006.
[2] B. Leimkuhler, S. Reich, *Simulating Hamiltonian Dynamics*. Cambridge Univ. Press, 2004.
[3] R.M. Wald, *General Relativity*. Univ. Chicago Press, 1984.
[4] A. Anderson, JW York Jr, *Hamiltonian Time Evolution for General Relativity*. Phys. Rev. Lett. 81, 1154–1157 (1998)

Solving Partial Differential-Algebraic Equations in Structural Mechanics

Christoph Lunk
(joint work with Bernd Simeon)

In my talk I present an algorithm for solving the equations of constrained structural dynamics with adaptivity in time and space. Giving a short outline the basis for this will be the Rothe method. We set up the mathematical model and formulate it as a time-dependent saddle point problem. For this purpose we need a time integration scheme and error estimators for the descretization in space and time.

The application fields, we are interested in, are characterized by mechanical multibody systems where each subsystem is discretized by finite elements or other methods in space. State of the art solution schemes separate the unknowns in space and time, using the Finite Elemente Methods (FEM) on a fixed mesh to discretize the spatial part. Together with constraints this yields a system of DAEs of index 3.

The advantage of this approach, also called Method of Lines (MoL), is that good time integration methods exist based on index reduction and stabilized integration. The disadvantage, however, is that adaptivity in space is hardly possibly.

The counterpart of the MoL, also called Rothe method, discretizes the system with respect to time first and solves the obtained stationary problem afterwards [2]. The Newmark algorithm and its generalizations are still the methods of choice for

time integration in structural dynamics. In the present work we shall show how the Newmark and generalized α-methods can be extended to differential-algebraic equations by using position and velocity stabilization as well as related techniques from molecular dynamics solvers as key ideas. The benefits of the α-methods are their controllable numerical dissipation properties. Thus spurious high frequent oscillations caused by the spatial discretization can be damped out effectively.

The behavior of elastic bodies is described by physical laws, e.g. by Cauchy's first law of motion

$$\rho\ddot{u}(x,t) = \mathrm{div}\sigma(u(x,t)) + \beta(x,t), \tag{1}$$

where $\sigma(u(x,t))$ is the stress tensor of the displacement $u(x,t)$, ρ is the mass density and $\beta(x,t)$ is a volume force. We generalize these kind of equations by writing

$$\ddot{u} = \mathcal{A}u + l.$$

Here $\mathcal{A}$ is an spatial operator, for instance $(1/\rho)\mathrm{div}\sigma(u)$ from (1), that can be nonlinear. The variable l denotes external loads. We skip the noncritical possibility that the right hand side depends on the velocity $\dot{u}$.

When we think of multibody systems, the variable u could contain the displacement of multiple subsystems. We express any coupling between these subsystems or other constraints with

$$\mathcal{B}u = m,$$

where $\mathcal{B}$ is an operator, e.g. the trace functional, which extracts the displacement of two bodies in contact.

Summarizing the saddle point formulation of the constrained movement reads

$$\begin{aligned} \ddot{u} - \mathcal{A}u + \mathcal{B}'\lambda &= l, \\ \mathcal{B}u &= m, \end{aligned} \tag{2}$$

where $\mathcal{B}'$ is the adjoint operator of $\mathcal{B}$ and λ the Lagrange mutliplier.

The usual way to solve this system is to introduce the velocity $w := \dot{u}$ and to form the first order index 3 equation system. Due to the unfavorable sensitivity of index 3 systems one differentiates the constraints with respect to time once

$$\mathcal{B}w = -\dot{\mathcal{B}}u + \dot{m}$$

and gets with (2) an index 2 formulation.

If we apply the Newmark scheme to the ordinary part of the equation and append the constraint forces, we achieve ($u_n \doteq u(\cdot,t_n), w_n \doteq w(\cdot,t_n)$)

$$\begin{aligned} \frac{w_{n+1}-w_n}{\triangle t} &= (1-\gamma)z_n + \gamma z_{n+1} - \frac{1}{2}\mathcal{B}'_n\lambda_n - \frac{1}{2}\mathcal{B}'_{n+1}\lambda_{n+1}, \\ \frac{u_{n+1}-u_n}{\triangle t} &= w_n + \triangle t(\frac{1}{2}-\beta)z_n + \triangle t\beta z_{n+1} - \triangle t(\frac{1}{2}-\bar{\beta})\mathcal{B}'_n\lambda_n - \triangle t\bar{\beta}\mathcal{B}'_{n+1}\lambda_{n+1}, \\ \mathcal{B}u_{n+1} &= m_{n+1}, \\ \mathcal{B}w_{n+1} &= -\dot{\mathcal{B}}u_{n+1} + \dot{m}_{n+1}. \end{aligned} \tag{3}$$

where $z_n = \mathcal{A}u_n + l_n$. The key idea of α-methods is the use of a convex combination

$$\alpha_m z_n + (1-\alpha_m) z_{n+1} = \alpha_f(\mathcal{A}u_n + l_n) + (1-\alpha_f)(\mathcal{A}u_{n+1} + l_{n+1}).$$

Here $\beta, \gamma, \alpha_f, \alpha_m$ and $\bar{\beta}$ are parameters of the time integration scheme. If one prescribes the spectral radius of numerical dissipation ρ_∞, i.e. the damping rate of eigenfrequencies $\omega \to \infty$, the suitable choice of parameters is $\alpha_f = \rho_\infty/(1+\rho_\infty), \alpha_m = (2\rho_\infty - 1)/(1+\rho_\infty)$. The scheme damps uniformly, iff $\beta = 1/4(1-\alpha_m+\alpha_f)^2$ and is of second order, iff $\gamma = 1/2 - \alpha_m + \alpha_f$. In this case the truncation error of u_{n+1} is $\bar{\beta}\triangle t^3 \dddot{u}_n + O(\triangle t^4)$ iff $\bar{\beta} = \beta + (\alpha_m - \alpha_f)/2$.

A special choice of parameters, namely $\alpha_m = \alpha_f, \gamma = 1/2, \beta = \bar{\beta} = 0$ leads directly to the RATTLE-scheme, which was developed for molecular dynamics. The combination of the RATTLE- and the α-schemes assigns the name *α-RATTLE scheme* for (3). Its convergence proof can be found in [3].

The first challenge is the estimation of the time integration error. Taylor expansion shows, that with

$$\tilde{u}_{n+1} := u_n + \frac{2\triangle t}{3}\dot{u}_n + \frac{\triangle t}{3}\dot{u}_{n+1} + \frac{\triangle t^2}{6}(\mathcal{A}u_n + l_n - \mathcal{B}_n'\lambda(\cdot, t_n))$$

we can control the error of u_{n+1} by minimization of the truncation error $\tilde{u}_{n+1} - u_{n+1} = \triangle t^3(\frac{1}{6} - \bar{\beta})\dddot{u}_n + O(\triangle t^4)$. We remark that the implicit given Lagrange multiplier λ_n and λ_{n+1} do not match the exakt values in (3). Indeed $\lambda_n = \lambda(\cdot, t_n) + \triangle t(1-2\bar{\beta})\dot{\lambda}(\cdot, t_n) + O(\triangle t^2)$ and $\lambda_{n+1} = \lambda(\cdot, t_n) + \triangle t 2\bar{\beta}\dot{\lambda}(\cdot, t_n) + O(\triangle t^2)$ hold. In case of a L-stable scheme it is $\bar{\beta} = 1/2$, i.e. both values are of first order.

The stationary equation system (3) can be solved by the FEM. In that case each line will be premultiplied with a test function and the function spaces will be projected on finite subspaces (cf. [4]). The operators $\mathcal{A}$ and $\mathcal{B}$ itself define functionals, e.g. $a(u,v) := \langle \mathcal{A}u, v\rangle$ where v is a test function of u and $l(v) := \langle l, v\rangle, b(u,\theta) := \langle \theta, \mathcal{B}u\rangle$ respectively. Under certain circumstances $a(\cdot,\cdot)$ defines a norm with which the spatial error of the solution u_{n+1} of the stationary system (3) can be estimated.

Now the overall algorithm reads: First one has to initialize a spatial mesh $\mathcal{T}_0$ at $t = t_0$ with basis functions $\{\phi_i\}$ for u and w and $\{\pi_j\}$ for λ_n and λ_{n+1} respectively. Additionally a time step size $\triangle t$ has to be defined and the initial values $u(x,t_0), \dot{u}(x,t_0)$ has to be projected on the finite subspace (gives u_0, w_0). Starting with $n = 0$ the program steps are

(1) assemble matrices $M_{ij} = \langle \phi_i, \phi_j\rangle_{ij}, K_{ij} = a(\phi_i, \phi_j)_{ij}$, $G_{ij} = b(\phi_i, \pi_j)$, $f_i = l_{n+1}(\phi_i)$ on mesh $\mathcal{T}_n$,
(2) project u_n, w_n, z_n on mesh $\mathcal{T}_n$,
(3) solve saddle point problem (Eq. (3)) in finite dimension,
(4) compute 3^{rd} order value $\tilde{u}_{n+1}$,
(5) estimate spatial error err_x of u_{n+1} and $\tilde{u}_{n+1}$, refine $\mathcal{T}_n$,
(6) if $err_x > tol_x$ go to (1), otherwise:
(7) estimate time error err_t, if $err_t > tol_t$ decrease $\triangle t$, fit z_n, go to (2), otherwise:

(8) remove nodes in $\mathcal{T}_n$, $\mathcal{T}_{n+1} := \mathcal{T}_n$, compute z_{n+1}, $n \mapsto n+1$, propagate new $\triangle t$, go to (1).

We remark that step (5) is essential, because the time error estimator should not be influenced by the spatial discretization error. The second hint concerns the projection from old to new mesh: only if the projection matches the spatial mesh in each time step exactly, undesirable oscillations can be prevented.

Finally we want to mention that this approach was successfully tested on a multibody system "pantograph with catenary" [1]. The differences between a solution with a fixed fine spatial grid and small time step sizes and our adaptive solution were below the given tolerances but the adaptive solution needed significantly less computer effort. Nevertheless, the proof of global convergence in time *and* space is still in work.

References

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

[2] F. Bornemann, M. Schemann, *An Adaptive Rothe Method for the Wave Equation*, Comput. Visual Sci. **1** (1998), 137–144.

[3] Ch. Lunk, B. Simeon, *Solving constrained mechanical systems by the family of Newmark and α-methods*, Preprint Series "Numerische Mathematik" **20** (2006), TU München.

[4] B. Simeon, *Numerische Simulation gekoppelter Systeme von partiellen und differential-algebraischen Gleichungen in der Mehrkörperdynamik*, Fortschritt-Berichte VDI, Reihe 20, **325** (2000), Düsseldorf.

Projector Based DAE Analysis

Roswitha März

Most approaches to more general DAEs are supplied by reduction techniques implying differentiations and eliminations. The use of derivative arrays involves the restriction to problems satisfying high smoothness demands. However, we are wrong in supposing more information in derivatives of a function than in the function itself. The tractability index concept provides an alternative way without the use of derivative arrays. It relies on projector based decompositions into characteristic parts and linearization. It allows to handle equations

$$f((d(x(t),t))', x(t), t) = 0, \tag{1}$$

where $f(y,x,t) \in \Re^m, d(x,t) \in \Re^n,\ y \in \Re^n,\ x \in \mathfrak{D}_0, t \in \Im$, $\mathfrak{D}_0 \subseteq \Re^m$ open, connected, $\Im \subseteq \Re$ an interval, f and d are continuous together with their first partial derivatives f_y, f_x, d_x, d_t. Denote, for $(x^1, x, t) \in \Re^m \times \mathfrak{D}_0 \times \Im$, $D(x,t) = d_x(x,t)$, $A(x^1,x,t,) = f_y(D(x,t)x^1 + d_t(x,t), x, t)$, $B(x^1,x,t,) = f_x(...)$. The DAE (1) is supposed to have a *properly stated leading term*, that is, the decomposition

$$\Re^n = \ker A \oplus \operatorname{im} D \tag{2}$$

is given pointwise on $\Re^m \times \mathfrak{D}_0 \times \Im$, both subspaces have constant dimension, and they are spanned by C^1 basises. Benefits from those refined DAE models have been realized first in [1]. For further arguments we refer to [2,3].

Regular linear DAEs with tractability index μ are supplied by constant-rank conditions for certain matrix functions $G_i, i = 0, ..., \mu$ (cf.[4-6]). For instance, all linear DAEs having a well-defined regular strangeness index satisfy these rank conditions. Regular linear DAEs can be completely decoupled by so-called *fine and completely decoupling projectors* into the (uniquely determined) inherent explicit regular ODE (IERODE), and the algebraic part including the inherent differentiations. These decoupled systems look like standard canonical forms (SCF)(cf.[7]), but with coefficients given explicitly in terms of the original DAE and the projectors computed from them. The resulting nilpotent block and its powers have constant rank. This allows for rigorous input-output relations, detailled qualitative stability analysis, etc.[2,8,9]. Points where the rank conditions fail are considered to be critical ones. For a detailed discussion of those *critical points* we refer to [10,11].

Quasi-regular linear DAEs form a class of DAEs with essentially relaxed rank-conditions. Since the nullspaces of the matrix functions G_i now may change their dimension, we do instead with continuous subnullspaces [12]. Quasi-regular DAEs can be decoupled in the same way as regular ones. Not surprisingly, all DAEs being transformable into SCF (with arbitrary nilpotent block) are quasi-regular. Now, rank changes of the nilpotent block are no longer indicated as critical points. However, note that this makes sense only for sufficiently smooth data oriented somehow on a "highest index subinterval". To be more precise we consider an example: The special DAE

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & \alpha(t) \\ 0 & 0 & 0 \end{pmatrix} (Dx(t))' + x(t) = q(t), \quad t \in [-1, 1], \tag{3}$$

with $\alpha(t) = t^{1/3}$ on $(0, 1]$, $\alpha(t) = 0$ on $[-1, 0]$, and $D = \mathrm{diag}(1, 0, 1)$, is quasi-regular. For each $q \in C, q_3 \in C^1$, there are C^1_D solutions. The restriction of this DAE to the subinterval $(0, 1]$ represents a regular DAE with tractability index two being also solvable for all $q \in C$,$q_3 \in C^1$. However, considering the restriction to the subinterval $[-1, 0]$, we may put $D = \mathrm{diag}(1, 0, 0)$, and a regular index-one DAE results that is solvable for all continuous q. Letting e.g. $q(t) = t^{1/3}$ we obtain at least continuous solutions on both subintervals, but the solution segments cannot be continuously glued to each other. Hence, the point $t_\star = 0$, which indicates a rank change in $G_0 = AD$, is rather a critical point. We call those critical points where nothing happens in a smoother setting *harmless critical points.*

Regular nonlinear DAEs with tractability index μ should be those the linearizations of which are regular with the same index. A linearization of (1) is

$$A_\star(t)(D_\star(t)x(t))' + B_\star(t)x(t) = q(t), \qquad t \in \Im_\star, \tag{4}$$

with $A_\star(t)=f_y((d(x_\star(t), t))', x_\star(t), t)$, $B_\star(t)=f_x(...)$, $D_\star(t)=D(x_\star(t), t)$, $t \in \Im_\star$, and $x_\star$ is a sufficiently smooth function on $\Im_\star \subseteq \Im$ with values in $\mathfrak{D}_0$. We will

realize this idea in a constructive way by means of matrix functions and projectors computed from the data in the original DAE (1). Choose $Q_0 \in C(\mathfrak{D}_0 \times \Im, L(\Re^m))$ to project pointwise onto $N_0(x,t) = \ker D(x,t)$, $P_0 = I - Q_0$, and determine the generalized inverse D^- by the relations $DD^-D = D, D^-DD^- = D^-, D^-D = P_0, DD^- = R$, where R denotes the continuously differentiable projector function onto $\operatorname{im} D$ along $\ker A$ (cf.(2)). Starting with $G_0 = AD, B_0 = B, \pi_0 = P_0$, we form, for $i \geq 0$,as long as the expressions exist, the matrix functions

$$(5) \qquad G_{i+1} = G_i + B_iQ_i, \quad B_{i+1} = B_iP_i - G_{i+1}D^-(D\pi_{i+1}D^-)'D\pi_i.$$

Thereby, Q_{i+1} denotes a projector onto $N_{i+1} = \ker G_{i+1}$, $P_{i+1} = I - Q_{i+1}, \pi_{i+1} = \pi_iP_{i+1}$. These relations are meant pointwise for $x \in \mathfrak{D}_0, t \in \Im$, and jet variables $x^j \in \Re^m$. The derivative involved in the expression for B_{i+1} is defined to be the total derivative in jet variables. Hence, in each level, a new variable x^{i+2} comes in, so that B_{i+1} depends on $x^{i+2}, ...x^1, x, t$.

The *projectors* $Q_0, ..., Q_\kappa$ in (5) *are said to be admissible on* $\mathfrak{D}_0 \times \Im$, if the relations

$$(6) \quad X_i \subseteq \ker Q_i, \quad X_i \oplus [(N_0 + ... + N_{i-1}) \cap N_i] = N_0 + ... + N_{i-1}, \quad i = 1, ..., \kappa,$$

become valid, and, for $i = 1, ..., \kappa$, Q_i is continuous on $\Re^{mi} \times \mathfrak{D}_0 \times \Im$, but $D\pi_iD^-$ is there continuously differentiable. By admissible projectors $Q_0, ..., Q_\kappa$ we obtain continuous $G_0, ..., G_{\kappa+1}$, and G_i has constant rank r_i, $i = 1, ..., \kappa$.

The *nonlinear DAE (1) is said to be regular on* $\mathfrak{D}_0 \times \Im$ *with tractability index* μ, if there are admissible on $\mathfrak{D}_0 \times \Im$ projectors such that $r_{\mu-1} < r_\mu = m$. The values $r_0, ..., r_\mu$ are called *characteristic values* of the DAE [12-14].

If a nonlinear DAE has Hessenberg-size-μ form (cf.[7]), then it is regular with tractability index μ and characteristic values $r_0 = ... = r_{\mu-1}$, and $m - r_{\mu-1}$ is the number of derivative-free equations [12].

The following assertion shows that we have realized in fact the above idea on an index notion via linearizations [12-14]: If the DAE (1) is regular on $\mathfrak{D}_0 \times \Im$ with tractability index μ and characteristic values $r_0, ..., r_\mu$, then all linearizations (4) are regular with the same index and characteristic values. Conversely, supposed some additional smoothness concerning the data of (1) is given, if all linearizations (4) are regular DAEs, then they have an uniform index and uniform characteristic values, and the nonlinear DAE (1) is regular on $\mathfrak{D}_0 \times \Im$ with that index and these characteristic values.

Our index notion allows for a priori and a posteriori index monitorings as e.g.discussed in [15]. We believe that those *diagnostic tools* will be in great demand just in the next future (cf.[14]). In a next step, a further localization of the index notion that concerns also the jet variables should be considered.

Extremal conditions for minimization problems with constraints described by DAEs are discussed in [16]. In particular, by a rigorous proof of a necessary optimality condition, a question left open for several years is now answered. This analysis relies on a deep projector based insight into the structure of regular DAEs and their adjoints [16,17].

References

[1] K.BALLA, and R.MÄRZ, *A unified approach to linear differential algebraic equations and their adjoints*,Z.Anal.Anw.21(2002)3,738-802.

[2] R.MÄRZ, *Differential algebraic systems anew*,Applied Numerical Mathematics 42(2002)315-335.

[3] I.HIGUERAS, and R.MÄRZ, *Differential algebraic equations with properly stated leading terms*,Computers and Mathematics with Applications 48(2004)215-235.

[4] R.MÄRZ, *The index of linear differential algebraic equations with properly stated leading terms*, Results in Mathematics 42(2002)308-338.

[5] R.MÄRZ, *Solvability of linear differential algebraic equations with properly stated leading term*, Results in Mathematics 45(2004)88-105.

[6] R.MÄRZ, *Fine decouplings of regular differential algebraic equations*, Results in Mathematics 46(2004)57-72.

[7] K.E.BRENAN, S.L.CAMPBELL,and L.R.PETZOLD, *Numerical solution of initial value problems in differential algebraic equations*,North-Holland Amsterdam, 1989.

[8] I.HIGUERAS, R.MÄRZ,and C:TISCHENDORF, *Stability preserving integration of index-1 DAEs*, Applied Numerical Mathematics 45(2003)175-200.

[9] I.HIGUERAS, R.MÄRZ,and C.TISCHENDORF, *Stability preserving integration of index-2 DAEs*, Applied Numerical Mathematics 45(2003)201-229.

[10] R.MÄRZ, and R.RIAZA, *Linear differential-algebraic equations with properly stated leading term:Regular points*, JMAA,to appear.

[11] R.MÄRZ, and R.RIAZA, *On linear differential algebraic equations with properly stated leading term:Critical points*, preprint 23-04, http://www.mathematik.hu-berlin.de/publ/pre/2004/P-04-23.ps.

[12] R.LAMOUR, R.MÄRZ, and C.TISCHENDORF, *Projector based DAE analysis*, manuscript in preparation.

[13] R.MÄRZ, *Characterizing differential algebraic equations without the use of derivative arrays*, Computers and Mathematics with Applications 50(2005)1141-1156.

[14] R.MÄRZ, *Differential algebraic systems with properly stated leading term and MNA equations*, Intern.Series of Numerical Mathematics 146(2003)135-151.

[15] R.LAMOUR, *Index determination and calculation of consistent initial values for DAEs*, Computers and Mathematics with Applications 50(2005)1125-1140.

[16] A.BACKES. *Extremalbedingungen für Optimierungsprobleme mit Algebro- Differentialgleichungen*, Dissertation,Humboldt-Universität zu Berlin, 2006.

[17] K.BALLA, *Differential algebraic equations and their adjoints*, Dissertation, Hungarian Academy of Sciences, Budapest, 2004.

Transformation of high order linear differential-algebraic systems to first order

Volker Mehrmann

(joint work with Chunchao Shi)

We study general linear l-th order systems of Differential-Algebraic Equations (DAEs) with variable coefficients

$$A_l(t)x^{(l)}(t) + A_{l-1}(t)x^{(l-1)}(t) + \cdots + A_0(t)x(t) = f(t), \tag{1}$$

in a real interval $\mathbb{I} \subset \mathbb{R}$, together with initial conditions

$$x(t_0) = x_0^{[0]},\ \ldots,\ x^{(l-2)}(t_0) = x_0^{[l-2]},\ x^{(l-1)}(t_0) = x_0^{[l-1]},\ t_0 \in \mathbb{I}. \tag{2}$$

Here, the coefficients satisfy $A_i(t) \in \mathcal{C}(\mathbb{I}, \mathbb{C}^{m,n})$, $i = 0, 1, \ldots, l$, $A_l(t) \not\equiv 0$, $x(t)$ is an unknown vector-valued function in $\mathcal{C}(\mathbb{I}, \mathbb{C}^n)$, and the right-hand side $f(t)$ is a given vector-valued function in $\mathcal{C}^k(\mathbb{I}, \mathbb{C}^m)$, where $\mathcal{C}^k(\mathbb{I}, \mathbb{C}^{m,n})$, $k \in \mathbb{N}_0$, denotes the set of all k-times continuously differentiable matrix-valued functions from the real interval $\mathbb{I}$ to the complex vector space $\mathbb{C}^{m,n}$ and k is sufficiently large. In the following we will refer to DAEs with order l greater than 1 simply as *high order* systems.

DAEs play a key role in the modeling and simulation of constrained dynamical systems in many applications. Such systems have been intensively studied, theoretically as well as numerically, in the past three decades. For a systematic and comprehensive exposition of important aspects regarding the theory, the numerical treatment and many applications of first order DAEs, see e.g. [2, 4, 8, 9, 13, 18] and the references therein.

Linear high order DAEs arise from linearizations of general nonlinear high order DAEs of the form

$$F(t, x, \dot{x}, \ldots, x^{(l)}) = 0 \tag{3}$$

arround reference solutions. Typical applications where second order DAEs arise naturally are multi-body systems, see [4, 18] or models of electrical circuits [6, 7].

Usually, in the classical theory of ordinary differential equations, high order systems are turned into first order systems by introducing new variables for the derivatives up to order $l-1$. There is no unique way of performing this transformation, and only recently for the case of constant coefficients (in the representation of matrix polynomials) a systematic theory for transformation to first order has been derived [15]. It has been indicated there, but also in several other publications, see [1, 3, 19], that the classical textbook approach of turning high order systems into first order form has to be performed with great care, since it may lead to substantial mathematical difficulties, in particular for DAEs.

We present the analysis of linear systems of differential-algebraic equations of higher order. This includes condensed forms for tuples of matrices and tuples of matrix-valued functions which are associated with the systems of constant and variable coefficients, respectively. Based on the condensed forms, we may convert such a system into an equivalent system, from which the behavior with respect to solvability, uniqueness of solutions and consistency of initial conditions can be directly read off.

We demonstrate that if one turns a higher order problem in the traditional way into a first order system of DAEs, then, to get the solvability and uniqueness of solutions, more smoothness of the right-hand side $f(t)$ may be required. The condensed forms, however, allow to do the transformation to first order without extra smoothness requirements.

Several issues remain open. These include the perturbation theory for higher order systems of DAEs, (see [10] for recent results in the case of constant coefficient systems) in particular how the decision making in the condensed forms influences the transformation to first order (see [16] for first results) as well as the construction

of appropriate numerical methods for the treatment of high order, high index differential-algebraic systems, see [19, 20] for first results.

References

[1] C. Arévalo and P. Lötstedt. Improving the accuracy of bdf methods for index 3 differential-algebraic equations. *BIT*, 35:297–308, 1995.

[2] K. E. Brenan, S. L. Campbell, and L. R. Petzold. *Numerical Solution of Initial-Value Problems in Differential Algebraic Equations*, volume 14 of *Classics in Applied Mathematics*. SIAM, Philadelphia, PA, second edition, 1996.

[3] C. De Boor and H. O. Kreiss. On the condition of the linear systems associated with discretized BVPs of ODEs. *SIAM J. Numer. Anal.*, 23:936–939, 1986.

[4] E. Eich-Soellner and C. Führer. *Numerical Methods in Multibody Systems*. B. G. Teubner Stuttgart, 1998.

[5] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, New York, 1982.

[6] M. Günther and U. Feldmann. CAD-based electric-circuit modeling in industry I. Mathematical structure and index of network equations. *Surv. Math. Ind.*, 8:97–129, 1999.

[7] M. Günther and U. Feldmann. CAD-based electric-circuit modeling in industry II. Impact of circuit configurations and parameters. *Surv. Math. Ind.*, 8:131–157, 1999.

[8] E. Hairer, C. Lubich, and M. Roche. *The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods*. Lecture Notes in Mathematics No. 1409. Springer-Verlag, Berlin, 1989.

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

[10] N. J. Higham, D. S. Mackey, and F. Tisseur. The conditioning of linearizations of matrix polynomials. Numerical Analysis Report No. 465, The University of Manchester, School of Mathematics, 2005, to appear in SIAM J. Matrix Analysis.

[11] P. Kunkel and V. Mehrmann. Canonical forms for linear differential-algebraic equations with variable coefficients. *J. Comput. Appl. Math.*, 56:225–259, 1994.

[12] P. Kunkel and V. Mehrmann. A new look at pencils of matrix valued functions. *Linear Algebra Appl.*, 212/213:215–248, 1994.

[13] P. Kunkel and V. Mehrmann. *Differential-Algebraic Equations — Analysis and Numerical Solution*. EMS Publishing House, Zürich, 2006.

[14] P. Kunkel, V. Mehrmann, und S. Seidel A MATLAB Toolbox for the Numerical Solution of Differential-Algebraic Equations. Preprint 16/2005, Institut für Mathematik, TU Berlin, 2002. `http://www.math.tu-berlin.de/preprints/`

[15] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Vector spaces of linearizations for matrix polynomials. Preprint 238, DFG Research Center Matheon, *Mathematics for key technologies* in Berlin, TU Berlin, Str. des 17. Juni 136, D-10623 Berlin, Germany, 2005. `http://www.matheon.de/,` to appear in SIAM J. Matrix Analysis.

[16] R. M. M. Mattheij and P. M. E. J. Wijckmans. Sensitivity of solutions of linear DAE to perturbations of the system matrices *Numer. Alg.*, 19:159–171, 1998.

[17] V. Mehrmann and C. Shi. Analysis of higher order linear differential-algebraic systems. Preprint 2004/17, Institut für Mathematik, TU Berlin, D-10623 Berlin, FRG, 2004. `url: http://www.math.tu-berlin.de/preprints/`.

[18] P. J. Rabier and W. C. Rheinboldt. *Nonholonomic Motion of Rigid Mechanical Systems from a DAE Viewpoint*. SIAM, Philadelphia, PA 19104-2688, USA, 2000.

[19] J. Sand. On implicit Euler for high-order high-index DAEs. *Appl. Numer. Math.*, 42:411–424, 2002.

[20] L. Wunderlich. *Numerical Solution of Second Order Differential-Algebraic Equations*. Diplomarbeit, TU Berlin. Inst. f. Mathematik, Berlin, FRG, 2004.

Runge-Kutta-Chebyshev Projection Method

Linda Petzold
(joint work with Zheming Zheng)

Projection methods have been widely used in the solution of incompressible Navier-Stokes equations, written in the nondimensionalized form:

$$\frac{\partial \mathbf{u}}{\partial t} + \nabla P = -(\mathbf{u}\cdot\nabla)\mathbf{u} + \frac{1}{Re}\nabla^2\mathbf{u}, \tag{1a}$$

$$\nabla \cdot \mathbf{u} = 0, \tag{1b}$$

with boundary conditions

$$\mathbf{u}|_\Gamma = \mathbf{u}_b, \tag{2}$$

where $\mathbf{u}$ is the velocity, P is the pressure and Re is the Reynolds number.

To solve this problem, projection methods use a fractional step approach, in which an intermediate velocity is obtained by solving the momentum equation (1a) without regard to the incompressibility constraint (1b), and then a projection of the intermediate velocity onto the divergence-free space is performed to obtain the corrected velocity that satisfies the incompressibility constraint. The pressure is computed in the projection step. In solving the incompressible Navier-Stokes equations with projection methods, much of the difficulty lies in the pressure update. The pressure does not evolve according to a differential equation. Rather, its value is determined by enforcing the incompressibility constraint. It has been observed that while the velocity can be reliably computed to second order accuracy in time, the pressure is typically only first order accurate in time.

If the incompressible Navier-Stokes equations are semi-discretized in space, they become a differential-algebraic equation (DAE) system. The mathematical structure of this DAE system is referred to as Hessenberg index 2. In the DAE context, $\mathbf{u}$ is the differential variable and P is the algebraic variable. The pressure P is further determined to be index 2, where the number of differentiations needed to determine the time derivative of P as a function of $\mathbf{u}$, P and t, is called the index of the DAE.

Often the momentum equation (1a) is solved implicitly in projection methods, due to the stiffness introduced by the viscous term. However we are motivated to develop an explicit projection method by solving the momentum equation explicitly with the use of a special purpose explicit Runge-Kutta method, the Runge-Kutta-Chebyshev (RKC) method [2], due to its enhanced stability properties. The RKC method was first proposed by Van der Houwen and Sommeijer. It was designed for the solution of moderately stiff ordinary differential equation (ODE) systems. This method exploits some remarkable properties of a family of explicit Runge-Kutta formulas of the Chebyshev type. This Runge-Kutta method uses the first two stages to achieve second order accuracy. The remainder of the stages are used to enlarge the stability region. It has the property of being stable while retaining a good accuracy using a minimum number of stages, and has been used

in the solution of parabolic partial differential equations discretized by the method of lines.

The explicit projection method we propose in this paper is called the Runge-Kutta-Chebyshev Projection (RKCP) method. In the RKCP method, the momentum equation is solved explicitly by the RKC method. One projection per step, regardless of the number of stages used in the RKC method, is performed at the last stage of the RKC method. An additional projection on the time derivative of the velocity, i.e., the acceleration, is performed to recover the second order temporal accuracy of the pressure, when it is desired. Because the RKC method was designed for the solution of moderately stiff ODE systems, the RKCP method is particularly well-suited for viscous dominated flows.

Our results are described in further detail in [1].

References

[1] Z. Zheng and L. Petzold, *Runge-Kutta Chebyshev Projection Method*, submitted, J. Comp. Phys. 2006.

[2] J. G. Verwer and B. P. Sommeijer, *An implicit-explicit Runge-Kutta-Chebyshev scheme for diffusion-reaction equations*, SIAM J. Sci. Comput. **25** (2004), 1824–.

Numerical Simulation of DAEs with Multiscale Behaviour in Time

Roland Pulch

The numerical simulation of electric circuits is based on a network approach, which yields systems of differential algebraic equations (DAEs). We write these systems in the general form

$$\frac{\mathrm{d}\mathbf{q}(\mathbf{x})}{\mathrm{d}t} = \mathbf{f}\left(\mathbf{b}(t), \mathbf{x}(t)\right) \qquad (\mathbf{x}, \mathbf{b} : \mathrm{R} \to \mathrm{R}^k,\ \mathbf{q}, \mathbf{f} : \mathrm{R}^k \to \mathrm{R}^k), \tag{1}$$

where $\mathbf{x}$ denotes unknown node voltages and branch currents. The function $\mathbf{b}$ represents predetermined input signals. In radio frequency applications, circuits exhibit amplitude and/or frequency modulated signals with widely separated time rates. Thus a transient integration of the corresponding system (1) becomes inefficient, since the fastest time scale restricts the size of time steps, whereas the slowest time scale determines the length of the time interval in the simulation. A multidimensional model for such signals yields an alternative approach by decoupling the time behaviour. Each separated time scale is given its own variable, which produces a multivariate function (MVF) of the signal. For amplitude modulated signals, Brachtendorf et al. [1] transformed the DAEs (1) into a system of multirate partial differential algebraic equations (MPDAEs). In case of two time scales, the system reads

$$\frac{\partial \mathbf{q}(\hat{\mathbf{x}})}{\partial t_1} + \frac{\partial \mathbf{q}(\hat{\mathbf{x}})}{\partial t_2} = \mathbf{f}\left(\hat{\mathbf{b}}(t_1, t_2), \hat{\mathbf{x}}(t_1, t_2)\right) \quad (\hat{\mathbf{x}}, \hat{\mathbf{b}} : \mathrm{R}^2 \to \mathrm{R}^k), \tag{2}$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{b}}$ represent the MVFs of $\mathbf{x}$ and $\mathbf{b}$, respectively. A solution of the MPDAE model (2) yields a solution of the original DAE (1) via $\mathbf{x}(t) = \hat{\mathbf{x}}(t, t)$.

Since the time scales are decoupled, the MVF can be computed using a relatively low number of grid points in time domain, i.e., an efficient numerical simulation is achieved. The determination of quasiperiodic signals in the DAEs (1) demands to solve the biperiodic boundary value problem

$$\hat{\mathbf{x}}(t_1,t_2)=\hat{\mathbf{x}}(t_1+T_1,t_2),\quad \hat{\mathbf{x}}(t_1,t_2)=\hat{\mathbf{x}}(t_1,t_2+T_2)\quad \text{for all } t_1,\ t_2\in\mathrm{R}. \tag{3}$$

If the fast time scale t_2 is periodic but the slow time scale t_1 is aperiodic, then initial-boundary value problems

$$\hat{\mathbf{x}}(0,t_2)=\mathbf{w}(t_2),\quad \hat{\mathbf{x}}(t_1,t_2)=\hat{\mathbf{x}}(t_1,t_2+T_2)\quad \text{for all } t_1\geq 0,\ t_2\in\mathrm{R} \tag{4}$$

with a predetermined function $\mathbf{w}$ arise. The selection of $\mathbf{w}(0)$ allows to reproduce intial value problems of the DAEs (1).

If the signals include amplitude as well as frequency modulation, then an additional time-dependent local frequency function is necessary to obtain an efficient multivariate representation. Narayan and Roychowdhury [3] introduced the corresponding system of warped MPDAEs

$$\frac{\partial \mathbf{q}(\hat{\mathbf{x}})}{\partial t_1}+\nu(t_1)\frac{\partial \mathbf{q}(\hat{\mathbf{x}})}{\partial t_2}=\mathbf{f}\left(\mathbf{b}(t_1),\hat{\mathbf{x}}(t_1,t_2)\right)\quad (\nu:\mathrm{R}\to\mathrm{R}). \tag{5}$$

Thereby, the local frequency function ν is unknown a priori, too. The input signals $\mathbf{b}$ act on the slow time scale t_1 only, whereas the system exhibits an inherent fast time scale. The reconstruction of a solution of the original DAEs (1) reads $\mathbf{x}(t)=\hat{\mathbf{x}}(t,\int_0^t \nu(s)\mathrm{d}s)$. Again biperiodic problems (3) or initial-boundary value problems (4) have to be considered. The determination of an adequate local frequency function is crucial for the efficiency of the multidimensional model, since inappropriate choices yield MVFs with undesired oscillations and thus require a fine grid in time domain.

In the previous work [4], the structure of the system of MPDAEs (5) was investigated in detail. Thereby, the system exhibits a transport of information along characteristic curves, i.e., a hyperbolic structure. Consequently, we constructed a method of characteristics in time domain, which solves biperiodic problems (3) efficiently in comparison to standard finite difference methods. A method of characteristics for solving initial-boundary value problems (4) is feasible but inefficient in case of widely separated time scales.

In the talk, we consider biperiodic boundary value problems (3) of the warped MPDAEs (5) again. We discuss the determination of an appropriate local frequency function here. Solutions of the warped system corresponding to different local frequency functions are interconnected by a transfomation formula. If $\hat{\mathbf{x}}$ and ν satisfy the system (5), then the transformed MVF

$$\hat{\mathbf{y}}(t_1,t_2):=\hat{\mathbf{x}}\left(t_1,t_2+\int_0^{t_1}\nu(s)-\mu(s)\,\mathrm{d}s\right) \tag{6}$$

represents a solution of (5) belonging to the given local frequeny function μ. To preserve the periodicities of the MVF, some restrictions on the choice of the local frequencies are necessary. Thus the structure implies that the local frequency

function represents free parameters in the multidimensional model. Narayan and Roychowdhury [3] proposed smooth phase conditions, which are used as additional boundary constraints, to identify a corresponding local frequency function. The strategy yields efficient representations by MVFs in general. However, this advantageous property can not be proved, since the approach via phase conditions is motivated just heuristically. We impose an optimisation criterion on the MVFs, which shall prevent undesired oscillations in the multidimensional representation, namely

$$\gamma\left(\hat{\mathbf{x}}\right) := T_1 \int_0^{T_1} \int_0^1 \sum_{l=1}^{k} w_l \left(\frac{\partial \hat{x}_l}{\partial t_1}\right)^2 \mathrm{d}t_2 \, \mathrm{d}t_1 \to \min. \quad \left(\hat{\mathbf{x}} = (\hat{x}_1, \ldots, \hat{x}_k)^\top\right) \tag{7}$$

with constant weights $w_l \geq 0$. A variational calculus based on the transformation (6) yields a necessary condition for an optimal solution, which reads

$$r(t_1) := \int_0^1 \sum_{l=1}^{k} w_l \cdot \frac{\partial^2 \hat{x}_l}{\partial {t_1}^2} \cdot \frac{\partial \hat{x}_l}{\partial t_2} \, \mathrm{d}t_2 = 0 \quad \text{for all } \; t_1 \in \mathrm{R}. \tag{8}$$

This additional condition can be included in a numerical scheme to determine the according optimal solution. Numerical simulations using a voltage controlled oscillator demonstrate that the constructed approach yields efficient representations by MVFs, where corresponding local frequencies are physically reasonable.

The periodicity in the slow time scale is crucial for the definition of the criterion (7) and the realisation of the corresponding variational calculus. A further field of research is the application of optimisation in case of initial-boundary value problems (4) for warped MPDAEs (5). Houben [2] already constructed a minimisation criterion, which is based on the charge term $\mathbf{q}(\hat{\mathbf{x}})$. This approach yields suitable solutions in general. On the other hand, alternative conditions based on the MVF $\hat{\mathbf{x}}$ itself shall be constructed to guarantee an appropriate minimisation directly. Smooth phase conditions can be applied in case of initial-boundary value problems, too. It is still an open question, if conditions based on minimisation can compete with phase conditions here.

References

[1] H. G. Brachtendorf; G. Welsch; R. Laur; A. Bunse-Gerstner: *Numerical steady state analysis of electronic circuits driven by multi-tone signals.* Electrical Engineering **79** (1996), 103–112.

[2] S.H.M.J. Houben: *Simulating multi-tone free-running oscillators with optimal sweep following.* In: W.H.A. Schilders; E.J.W. ter Maten; S.H.M.J. Houben (eds.): Scientific Computing in Electrical Engineering, Mathematics in Industry, Springer, 2004, 240–247.

[3] O. Narayan; J. Roychowdhury: *Analyzing oscillators using multitime PDEs.* IEEE Trans. CAS I **50** (2003) 7, 894–903.

[4] R. Pulch: *Multi time scale differential equations for simulating frequency modulated signals.* Appl. Numer. Math. **53** (2005) 2-4, 421–436.

Linear and Time-Invariant Abstract Differential-Algebraic Systems

Timo Reis

In today's engineering applications, there is an increasing interest in partial differential-algebraic equations (PDAEs), which are mainly coupled systems of partial differential equations (PDEs) and differential-algebraic equations (DAEs). This type appears e.g. in modeling of electrical circuits with further components which are modeled by PDEs. These can be parasitic like heat conduction or transmission lines [1, 8] as well as they could be the result of a more reliable modeling of complex components like semiconductor devices [2, 11, 12]. Moreover, PDAEs are the outcome of mathematical models of several mechanical systems like elastic multibody systems [4] or biomechanical systems like blood flow networks. In order to study these problems in a systematic way, we are led to differential-algebraic systems $F(\dot{x}(t), x(t), t) = 0$ in an abstract setting, the so-called abstract DAEs (ADAEs). The unknown function $x(\cdot)$ is now a path in an appropriate (mostly infinite dimensional) Hilbert space, and the Frechét derivative $\frac{\mathrm{d}}{\mathrm{d}\dot{x}}F(\dot{x}, x, t)$ has a nontrivial nullspace, in general. Here, we focus on the linear constant coefficient case

$$E\dot{x}(t) = Ax(t) + f(t). \tag{1}$$

$E : X \to Z$ is now a bounded linear operator and X, Z are some separable Hilbert spaces. In many practical cases, A is often acting on some product spaces and it is a block operator containing differential and evaluation operators. Hence, it is natural to assume that it is unbounded in general and that it is defined on some proper subspace $D(A) \subset X$.

In the finite dimensional version of (1), i.e. E and A are square matrices, the Kronecker normal form is a powerful theoretical tool for the analysis. In this case, a state space transform of (1) leads to the following decoupled differential equations

$$N\dot{x}_1(t) = x_1(t) + f_1(t) \tag{2a}$$

$$\dot{x}_2(t) = \bar{A}x_2(t) + f_2(t), \tag{2b}$$

where N is nilpotent and $\bar{A}$ is some square matrix. The nilpotency index $\nu \in \mathbb{N}$ of N is well-defined by the pair (E, A) and is called the *Kronecker index*. Based on this representation, the set of consistent initial values can be determined (see [3], for instance). x_{20} can be chosen arbitrarily whereas x_{10} has to satisfy $N^{k+1}x_1^{(k+1)}(0) = N^k x_1^{(k)}(0) + N^k f_1^{(k)}(0)$ for $k = 0, \ldots, \nu - 1$ that comes from a successive formal differentiation and multiplication from the left with N to (2a) and particularly considering $t = 0$. These relations are called the *algebraic* and *hidden algebraic constraints*. Further, the sensitivity of $x(t)$ with respect to derivatives of the inhomogeneity $f(t)$ can be measured leading to the notion of *perturbation index*. In [9], the question was treated whether it is possible to generalize the Kronecker form to the infinite dimensional case. Thereby, the concept

of *decoupling form* was developed. An ADAE in decoupling form is given by

$$\begin{aligned} \begin{pmatrix} N & 0 \\ 0 & I \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} &= \begin{pmatrix} I & K \\ 0 & \mathfrak{U} \\ 0 & \mathfrak{P} \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + \begin{pmatrix} f_1(t) \\ f_2(t) \\ f_3(t) \end{pmatrix} \\ \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} &= \begin{pmatrix} x_{10} \\ x_{20} \end{pmatrix}. \end{aligned} \tag{3}$$

The bounded operator N is nilpotent with nilpotency index ν, a number that is called *ADAE index*. This concept was first published in [6] as a generalization of the tractability index (see [7]) to infinite dimensions. The main differences to the finite dimensional case is the appearance of the operators $\mathfrak{P}$ and K. The operator $\mathfrak{P}$ appearing in the third row of the decoupling form has its interpretation as a boundary control term. The *coupling operator* K is not always eliminable in contrast to the finite dimensional case. In [9, 10], examples of ADAEs were given that do not possess a decoupling form with $K = 0$. [10] gives sufficient criteria form the removability of K.

The benefit of the decoupling form (3) for ADAEs is that the set of consistent initial values can be determined and perturbation results can be derived [10]. However, the appearance of the boundary and coupling operator leads to additional difficulties for the parameterization of the consistent initial values of ADAEs. Indeed, the initial value not only has to fulfill (hidden) algebraic constraints but also some further relations which can be obtained by formally differentiating the third row of (3).

The results are applied to ADAEs modeling electrical circuits with transmission lines. Based on the decoupling form (3) and the results of [5] for circuits with lumped elements, [10] gives circuit topological criteria for their index and consistent initial values.

References

[1] A. Bartel, *First order thermal PDAE models in electric circuit design*, Troch, I., F. Breitenecker, F. (eds.) Proc. 4th Mathmod, Vienna (2003).

[2] M. Bodestedt and C. Tischendorf, *PDAE Models of Integrated Circuits and Index Analysis*, To appear in "Math. Comput. Model. Dyn. Syst.".

[3] S.L. Campbell, *Singular Systems of Differential Equations*, Pitman Advanced Publishing Program (1980).

[4] J. Diaz and C. Führer, *A wavelet semidiscretisation of elastic multibody systems*, ZAMM **83** (2003), 677-689.

[5] D. Estévez Schwarz and C. Tischendorf, *Structural analysis of electrical circuits and consequences for the MNA*. International Journal of Circuit Theory and Applications, **28** (2000), 131-162.

[6] R. Lamour and R. Maerz and C. Tischendorf, *PDAEs and Further Mixed Systems as Abstract Differential Algebraic Systems* Preprint, Humboldt Universität zu Berlin (2001).

[7] R. Maerz, *Solvability of linear differential-algebraic equations with properly stated leading terms* , Result. Math. **45**, 88-105 (2002).

[8] T. Reis, *An Infinite Dimensional Descriptor System Model for Electrical Circuits with Transmission Lines*, Proc. MTNS 2004, Leuven, Belgium (2004).

[9] T. Reis and C. Tischendorf, *Frequency Domain Methods and Decoupling of Linear Infinite Dimensional Differential Algebraic Systems*, J. Evol. Equ., **5** (2005), 357-385.

[10] T. Reis, *Systems Theoretic Aspects of PDAEs and Applications to Electrical Circuits*, Doctoral Thesis, Technische Universität Kaiserslautern (2006).
[11] S. Schulz, *Ein PDAE-Netzwerkmodell als Abstraktes Differential-Algebraisches System*, Diploma Thesis, Humboldt Universität zu Berlin (2002).
[12] C. Tischendorf, *Coupled systems of differential-algebraic and partial differential equations in circuit and device simulation. Modeling and numerical analysis*, Habilitation Thesis, Humboldt Universität zu Berlin (2003).

Singularities of differential-algebraic equations

Ricardo Riaza
(joint work with Roswitha März)

Singular points of differential-algebraic equations (DAEs) can be roughly defined as those where the assumptions supporting an index notion fail. In the present contribution, we tackle singular (or, better, *critical*) points of linear DAEs with a properly stated leading term, that is, DAEs of the form

$$A(t)(D(t)x(t))' + B(t)x(t) = q(t), \quad t \in \mathcal{J}, \tag{1}$$

where $\mathcal{J} \subseteq \mathbb{R}$ is an interval, and the matrix coefficients $A(t) \in L(\mathbb{R}^n, \mathbb{R}^m)$, $D(t) \in L(\mathbb{R}^m, \mathbb{R}^n)$, $B(t) \in L(\mathbb{R}^m)$ depend continuously on t.

1. Regular problems: the P-framework

The leading term of the DAE (1) is said to be properly stated on the interval $\mathcal{J}$ if the coefficients $A(t)$ and $D(t)$ satisfy $\ker A(t) \oplus \operatorname{im} D(t) = \mathbb{R}^n$ for all $t \in \mathcal{J}$, and both subspaces have constant dimension and are spanned by C^1 basis functions. Assuming the leading term of (1) to be stated properly on the interval $\mathcal{J}$, denote as $R(t)$ the C^1 projector function realizing the decomposition above with $\operatorname{im} R(t) = \operatorname{im} D(t)$, $\ker R(t) = \ker A(t)$, $t \in \mathcal{J}$.

In the sequel we drop the argument t in the matrix functions involved. Introduce

$$G_0 := AD, \quad B_0 := B. \tag{2}$$

If the leading term is properly stated, then G_0 has constant rank r_0 on $\mathcal{J}$. Defining $N_0 := \ker D = \ker G_0$, let P_0 be any continuous projector along N_0, and take $Q_0 := I - P_0$. Additionally, denote as D^- the continuous on $\mathcal{J}$ generalized inverse of D uniquely defined by the four conditions

$$DD^-D = D, \quad D^-DD^- = D^-, \quad DD^- = R, \quad D^-D = P_0. \tag{3}$$

For $i \geq 1$, define

$$G_i := G_{i-1} + B_{i-1}Q_{i-1}. \tag{4}$$

If G_i has constant rank r_i, let $N_i := \ker G_i$, and choose a continuous projector P_i along N_i. Write

$$B_i := B_{i-1}P_{i-1} - G_iD^-(DP_0 \cdots P_iD^-)'DP_0 \cdots P_{i-1}. \tag{5}$$

The sequence is continued by defining $Q_i = I - P_i$, G_{i+1}, etc. A non-singular matrix function G_μ (with singular G_i, $i < \mu$) defines the system as *regular with tractability index* μ [1], and the solutions of problems with arbitrarily high index can be described via a decoupling of the DAE [2]. A local version of this regularity concept, supporting the notion of a *regular point*, is introduced in [3]: critical points are therefore defined as those not satisfying this regularity notion [4].

Note that in order to build the matrix chain (4)-(5), we make use of three assumptions in every step: (a) G_i has constant rank; additionally, the projectors Q_i are required to satisfy $Q_iQ_j = 0$, for all $0 \le j < i$: the existence of such a projector Q_i relies on the condition (b) $(N_0 \oplus \cdots \oplus N_{i-1}) \cap N_i = \{0\}$ on $\mathcal{J}$, what in turn supports writing the direct sum $N_0 \oplus \cdots \oplus N_{i-1} \oplus N_i$ in the next step. Finally, (c) the products $DP_0 \cdots P_iD^-$ are assumed to be C^1.

2. Critical points

As detailed below, for sufficiently smooth problems, the failing of conditions (a) and (b) above characterize the critical points of the DAE.

Theorem [4]**.** *Assume that the coefficients $A(t)$, $D(t)$, $B(t)$ in the DAE (1) are C^{m-1}. Then every critical point t_* of the DAE belongs to one of the following invariant, independent of projectors types:*

(i) *type* 0 *if G_0 has a rank drop at t_*;*
(ii) *type k-A, $k \ge 1$, if it is not type* 0*, j-A or j-B with $j < k$, and G_k has a rank drop at t_*;*
(iii) *type k-B, $k \ge 1$, if it is not type* 0*, j-A or j-B with $j < k$, nor k-A, and $N_k(t_*) \cap \{N_0(t_*) \oplus \cdots \oplus N_{k-1}(t_*)\} \ne \{0\}$.*

In order to define a working scenario accommodating A- and B-critical points, we construct below the tractability chain in a different, purely recursive manner.

3. The Π-framework

An alternative, simpler construction of the tractability matrix chain stems from the fact that not individual projectors P_i, Q_i but products of the form $P_0 \dots P_i$ and $P_0 \dots P_{i-1}Q_i$ are needed in the matrix chain construction, together with the property that $P_0 \dots P_i$ is along $K_i = N_0 \oplus \dots \oplus N_i$ (see e.g. [3, Proposition 1]).

The leading term of (1) is assumed to be properly stated. Define, as before, $G_0 := AD$. Using the constant rank of G_0 which follows from the proper statement of the DAE, choose a continuous projector Π_0 along $N_0 := \ker G_0$, and let $\Gamma_0 := I - \Pi_0$. Denote $K_0 := N_0$, let D^- be given by the conditions (3) with $P_0 = \Pi_0$, and define $C_0 := B$.

Now, for $i \ge 1$, define

$$G_i := G_{i-1} + C_{i-1}\Gamma_{i-1}. \tag{6}$$

If G_i is singular, denote $N_i := \ker G_i$ and check whether (a) G_i has constant rank, and also whether (b) $K_{i-1} \cap N_i = \{0\}$. If both conditions are met, proceed by

choosing a continuous projector

$$\Pi_i \text{ along } K_i := K_{i-1} \oplus N_i, \text{ with } \operatorname{im}\Pi_i \subseteq \operatorname{im}\Pi_{i-1}, \tag{7}$$

and define

$$\Gamma_i := \Pi_{i-1} - \Pi_i. \tag{8}$$

Finally, if (c) $D\Pi_i D^-$ is C^1, complete the i-th step by constructing

$$C_i := C_{i-1} - G_i D^- (D\Pi_i D^-)' D. \tag{9}$$

Note that the image condition in (7) can be met since $K_{i-1} \oplus \operatorname{im}\Pi_{i-1} = \mathbb{R}^m$ and hence there exists a space transversal to $K_{i-1} \oplus N_i$ within $\operatorname{im}\Pi_{i-1}$. It is also worth emphasizing that the image condition in (7) is satisfied automatically if Π_i is chosen as the *orthogonal* projector along $K_i = K_{i-1} \oplus N_i$. Provided that the smoothness condition (c) is met, this yields a well-defined criterion for the choice of projectors in the tractability index construction.

From the construction above it follows that a DAE is regular with index μ iff there exists a Π-sequence satisfying the requirements in (7) and for which G_μ is non-singular, with G_i singular if $i < \mu$. But this new framework provides a setting within which a broad class of critical problems can be handled, as shown below.

4. Scalarly implicit decoupling of critical DAEs

Theorem. *Assume that the set of regular points $\mathcal{J}_{\rm reg}$ is dense in $\mathcal{J}$, and that there exist projector functions*

1. *$R \in C^1(\mathcal{J}, L(\mathbb{R}^n))$ satisfying $\operatorname{im} R = \operatorname{im} D$ and $\ker R \subseteq \ker A$, for all $t \in \mathcal{J}$; and*
2. *$\Pi_1, \ldots, \Pi_{m-1}$, continuous on $\mathcal{J}$, with $D\Pi_i D^-$ continuously differentiable on $\mathcal{J}$, and satisfying $\operatorname{im}\Pi_i \subseteq \operatorname{im}\Pi_{i-1}$, such that, for $t \in \mathcal{J}_{\rm reg}$, $\ker \Pi_i = K_i$.*

Then, letting $\omega_\mu = \det G_\mu$, there exist continuous operators $\tilde{\mathcal{K}}_k$, $\tilde{\mathcal{L}}_k$, $\tilde{\mathcal{N}}_{kj}$, $\tilde{\mathcal{M}}_{kj}$ such that $x \in C^1_D(\mathcal{I}, \mathbb{R}^m) := \{x \in C(\mathcal{I}, \mathbb{R}^m) : Dx \in C^1(\mathcal{I}, \mathbb{R}^n)\}$ solves (1) in a given subinterval $\mathcal{I} \subseteq \mathcal{J}$ if and only if it can be written as

$$x = D^- u + v_0 + \cdots + v_{\mu-1}, \tag{10}$$

where $u \in C^1(\mathcal{I}, \mathbb{R}^n)$ is a solution of the scalarly implicit ODE

$$\omega_\mu u' - \omega_\mu (D\Pi_{\mu-1} D^-)' u + D\Pi_{\mu-1} G_\mu^{adj} B D^- u = D\Pi_{\mu-1} G_\mu^{adj} q,$$

on the locally invariant space $\operatorname{im} D\Pi_{\mu-1} D^-$, whereas the solution components $v_k \in C^1_D(\mathcal{I}, \mathbb{R}^m)$, $k = 1, \ldots, \mu-1$, $v_0 \in C(\mathcal{I}, \mathbb{R}^m)$ verify

$$\begin{aligned}
\omega_\mu v_{\mu-1} &= -\tilde{\mathcal{K}}_{\mu-1} D^- u + \tilde{\mathcal{L}}_{\mu-1} q, \\
\omega_\mu^{\mu-k} v_k &= -\tilde{\mathcal{K}}_k D^- u + \tilde{\mathcal{L}}_k q + \sum_{j=k+1}^{\mu-1} \tilde{\mathcal{N}}_{kj} (D v_j)' + \sum_{j=k+2}^{\mu-1} \tilde{\mathcal{M}}_{kj} v_j, \quad {\scriptstyle k=\mu-2,\ldots,1,0.}
\end{aligned}$$

The key assumptions in this theorem express the existence of a continuous extension of the "sum" spaces K_i preserving the transversality property depicted

in (b); it is worth emphasizing that these hypotheses hold for properly stated reformulations of standard-form linear DAEs with analytic coefficients.

References

[1] R. März, *The index of linear differential algebraic equations with properly stated leading terms,* Results in Mathematics **42** (2002) 308-338.
[2] R. März, *Solvability of linear differential algebraic equations with properly stated leading terms,* Results in Mathematics **45** (2004), 88-105.
[3] Roswitha März and Ricardo Riaza, *Linear differential-algebraic equations with properly stated leading term: Regular points,* J. Math. Anal. Appl., in press (2006).
[4] Roswitha März and Ricardo Riaza, *Linear differential-algebraic equations with properly stated leading term: A-critical points,* Math. Comput. Model. Dyn. Sys., in press (2006).

DAE's and Beyond: From Constrained Mechanical Systems to Saddle Point Problems

Bernd Simeon

Computational mechanics and its various applications in vehicle analysis, aerospace engineering, robotics, and materials science have experienced a significant development over the last decades. From the numerical analysis point of view, the question of constraints and their discretization is one of the key issues in this field.

In many cases, a dynamic saddle point problem is at the core of the problem formulation. We can write the saddle point problem as

$$\begin{aligned} \dot{u} + Au + B'\mu &= l \\ Bu &= m \end{aligned}$$

with $u(x,t)$ as primal unknown, e.g., the displacement or velocity field, and the Lagrange multiplier $\mu(x,t)$. The operator A is typically an elliptic operator while B stands for the constraints and B' for the corresponding dual operator.

After discretization in space, this infinite dimensional DAE or PDAE system turns into a linear, semi-explicit DAE of index two if the constraints have full rank. In this context, we notice the connection to saddle point theory and the inf-sup condition.

The first part of the talk at the Oberwolfach workshop on DAE's concentrated on this relationship between semi-explicit DAE's and PDE models in saddle point formulation, e.g., the Navier-Stokes equations in the incompressible case, domain decomposition approaches, and flexible multibody systems.

In the second part of the talk, materials with memory were investigated from a DAE perspective. Examples for this problem class are elastoplasticity and shape memory alloys, and the applications comprise the stretch formation of metal sheets and micromechanical devices. In general, the mathematical models consist of a coupled system of balance equations and unilaterally constrained evolution equations, Again, a saddle point formulation arises as natural problem setting and starting point for numerical analysis.

We presented a convergence result for implicit Runge-Kutta methods when applied to the infinite-dimensional system of constrained evolution equations Furthermore, the algorithms implemented in the FEM have been improved by state-of-the-art techniques available in DAE solvers, and recently, shape memory wires as actuators for robotics have been simulated successfully.

References

[1] Büttner, J., Simeon, B.: *Time Integration of the Dual Problem of Elastoplasticity by Runge-Kutta Methods.* SIAM J. Numerical Analysis 41(4), 1564-1584 (2003)

[2] Teichelmann, G., Helm, D., Simeon, B.: *Modelling and Simulation of Shape Memory Alloys.* In: I. Troch, F. Breitenecker (Ed.): 5th MATHMOD Vienna. Proceedings IMACS Symposium on Mathematical Modelling, Argesim Report No. 26, 2006

Grid Density Control: What should an adaptive ODE solver do?

Gustaf Söderlind

In recent year there has been a quick progress in the development of a mathematical theory for grid adaption. Today, adaptive time–stepping can be based on control theory and signal processing. A similar approach is suggested for two-point boundary value problems in order to create adaptive grids, based on controlling a grid density function. Combining adaptive time–stepping with grid density control opens up new possibilities for grid refinement and moving mesh algorithms. The present talk will discuss some basic techniques and open problems related to grid density control.

In discretization methods there is a trade-off between accuracy and computational effort. For efficiency, one wants as few grid points as possible, and try to put them where they really matter to accuracy. Many different approaches have been suggested in connection with both time–stepping and grid generation in boundary value problems. Here we shall develop adaptive grid generation based on controlling the grid density.

To this end we introduce a differentiable transformation, $x = \Gamma(\xi)$, where x is the original independent variable and ξ is a new, formal ("logical") independent variable. The transformation is supposed to satisfy $\Gamma'(\xi) = 1/\rho(\xi) > 0$. The condition $\rho(\xi) > 0$ implies that x is a monotone function of ξ.

The approach differs a little in initial value and boundary value problems. In the former case, we introduce an equidistant grid in ξ and denote its constant step size by ε. The sampling correspondence

$$x_{i+1} - x_i = \Gamma(\xi_{i+1}) - \Gamma(\xi_i) \approx \frac{\varepsilon}{\rho(\xi_{i+1/2})},$$

where $\xi_{k+1/2} = (\xi_{k+1} + \xi_k)/2$. This relates the uniform grid in ξ to a nonuniform grid in x, and the function ρ is interpreted as a *mesh density function.* The

transformation further allows us to rewrite the original differential equation

$$\dot{y} = f(y),$$

where dot denotes derivative with respect to x, by an augmented system,

$$y' = f(y)/\rho; \quad \rho' = g(y); \quad x' = 1/\rho.$$

Here prime is derivative with respect to ξ, and the added equations represent a control law for generating ρ and for recovering the original variable x, respectively.

The augmented system is discretized and solved using a constant step size method in ξ. This corresponds to using a variable step size method in x, but has the advantage that no special techniques are needed for the adaptivity, which is represented by the variation of ρ, while convergence is handled by letting $\varepsilon \to 0+$. In this way, it is technically possible to prove convergence of an adaptive method. The approach has also proved to be of great value in the simulation of Hamiltonian systems, where *near-conservation of energy is retained in spite of varying step size*. Examples will be given to demonstrate its performance.

In two-point boundary value problems, the technique takes a slightly different form. Here, we typically work with a fixed interval and a fixed number of grid points. Thus, for an arbitrary $N > 0$, let $\xi_k = k/(N+1)$, and introduce the equidistant grid $\Xi_N = \{\xi_k\}_{k=1}^N$, with $\bar{\Xi}_N = \{0, \Xi_N, 1\}$. As before we denote its mesh width by $\varepsilon_N = 1/(N+1)$. Further, suppose that we have a function Γ satisfying the conditions above. Then Γ deforms the grid Ξ_N via the map

$$X_N := \Gamma(\Xi_N) \quad \Leftrightarrow \quad x_k = \Gamma(\xi_k); \quad k = 1 : N,$$

which again implies that

$$x_{k+1} - x_k = \Gamma(\xi_{k+1}) - \Gamma(\xi_k) \approx \Gamma'(\xi_{k+1/2}) \cdot (\xi_{k+1} - \xi_k) = \frac{\varepsilon_N}{\rho(\xi_{k+1/2})}.$$

The grid X_N is non-uniform unless $\rho(\xi) \equiv 1$; it is "dense" or "fine" where ρ is large; and "sparse" or "coarse" where ρ is small.

Just as in the initial value problem case, obtaining an adaptive grid is a matter of finding a suitable function ρ. In the time-stepping case, ρ is successively generated along with the time integration, but in the boundary value case, a given ρ corresponds to a given grid. Therefore, the technique is rather used in the latter case for *grid refinement*.

We report preliminary results an a new grid refinement algorithm. It works by selecting a fairly small, fixed value of N, and then employs an Euler–Lagrange criterion for successively refining the grid, so that the grid points are located where they improve accuracy; the refinements typically go on for five to ten steps. After this step, when the function ρ has been found, the grid is *oversampled* from ρ, and the total number of grid points N is determined so that the error estimate approximately equals a given tolerance, for a given error criterion. Examples will be given from simple test runs.

Finally, as grid refinement for 2pBVP can be viewed as a procedure updating the grid in a "pseudo time," it is immediately seen that one can combine the adaptive time-stepping technique in true time, with grid refinement for 2pBVPs. This

opens up new possibilities for moving mesh algorithms in PDEs. The possible arrangement of such algorithms will be briefly discussed.

Supported by Swedish Research Council VR grant 2005xyz.

Control problems for differential-algebraic equations

Tatjana Stykel

In this report we briefly discuss stability, passivity and model order reduction of linear time-invariant control systems described by differential-algebraic equations (DAEs)

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \end{aligned} \tag{1}$$

where $E, A \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{n,m}$, $C \in \mathbb{R}^{p,n}$, $D \in \mathbb{R}^{p,m}$, $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^m$ is the input, and $y(t) \in \mathbb{R}^p$ is the output. Such equations arise in a variety of applications including multibody dynamics and circuit simulation.

It is well known that the stability properties of system (1) can be characterized in terms of the eigenvalues of the pencil $\lambda E - A$. System (1) with $u(t) \equiv 0$ is *asymptotically stable* if and only if $\lambda E - A$ is stable, i.e., all the finite eigenvalues of $\lambda E - A$ have negative real part. Note, however, that the eigenvalues of $\lambda E - A$ may be very ill-conditioned in the sense that they may change largely even for small perturbations in E and A. Hence, eigenvalues that are computed numerically in finite precision arithmetic, may not always provide the correct information on the stability of dynamical systems. As an alternative to the use of eigenvalues in the stability analysis, one can employ spectral parameters based on projected Lyapunov equations [5, 6]. One can show that the pencil $\lambda E - A$ is stable if and only if the projected generalized continuous-time Lyapunov equation

$$A^T H E + E^T H A = -P_r^T P_r, \qquad H = P_l^T H P_l \tag{2}$$

has a unique symmetric, positive semidefinite solution H. Here P_r and P_l are the spectral projectors onto the right and left deflating subspaces of the pencil $\lambda E - A$ corresponding to the finite eigenvalues. The parameter $\kappa(E, A) = 2\|E\|\|A\|\|H\|$, where $\|\cdot\|$ denotes the spectral matrix norm, can be used to characterize the stability of $\lambda E - A$ and also the sensitivity of its eigenvalues to perturbations in the matrices E and A, see [5].

Passivity is an important concept in circuit simulation. System (1) is *passive* if and only if its transfer function $\mathbf{G}(s) = C(sE - A)^{-1}B + D$ is positive real, i.e., $\mathbf{G}(s)$ is analytic in $\mathbb{C}^+ = \{\, s \in \mathbb{C} \,:\, \mathrm{Re}(s) > 0 \,\}$ and the matrix $\mathbf{G}(s) + \mathbf{G}^T(\bar{s})$ is positive semidefinite for all $s \in \mathbb{C}^+$. We have the following result.

Proposition. *Let $\mathbf{G}(s) = \mathbf{G}_{sp}(s) + \mathbf{P}(s)$, where $\mathbf{G}_{sp}(s)$ is the strictly proper part and $\mathbf{P}(s) = P_0 + sP_1 + \ldots + s^q P_q$ is the polynomial part of $\mathbf{G}(s)$.*

1. *If P_1 is symmetric, positive semidefinite, $P_j = 0$ for $j \geq 2$ and if the projected generalized Lur'e equation*

$$\begin{aligned} A^T\mathcal{Y}E + E^T\mathcal{Y}A &= -P_r^T L^T L P_r, \quad \mathcal{Y} = P_l^T \mathcal{Y} P_l, \\ B^T\mathcal{Y}E - CP_r &= -K^T L P_r, \quad K^T K = P_0 + P_0^T \end{aligned} \tag{3}$$

has the solution $\mathcal{Y}$, L, K, where $\mathcal{Y}$ is symmetric and positive semidefinite, then $\mathbf{G}(s)$ is positive real.

2. *If $\mathbf{G}(s)$ is positive real and if system* (1) *is minimal, then the projected generalized Lur'e equation* (3) *has the solution $\mathcal{Y}$, L and K.*

If $R = P_0 + P_0^T$ is nonsingular, then the projected Lur'e equation (3) is equivalent to the projected generalized Riccati equation

$$A^T\mathcal{Y}E + E^T\mathcal{Y}A + (B^T\mathcal{Y}E - CP_r)^T R^{-1}(B^T\mathcal{Y}E - CP_r) = 0, \quad \mathcal{Y} = P_l^T \mathcal{Y} P_l.$$

Modelling of complex physical and technical processes such as VLSI chip design and control of fluid flow often leads to linear DAE control systems of very large order n, while the number m of inputs and the number p of outputs are typically small compared to n. Despite the ever increasing computational speed, simulation, optimization or real-time controller design for such large-scale systems is difficult because of large storage requirements and computation time. In this context, *model order reduction* is of crucial importance. A general idea of model reduction is to approximate the large-scale system (1) by a reduced-order model that preserves essential properties of (1) like stability and passivity and that has a small approximation error.

Balanced truncation is one of the most effective and well studied model reduction approaches for standard state space systems [2, 4]. This approach has been extended to DAE systems in [7]. An important property of the balanced truncation model reduction methods is that the asymptotic stability is preserved in the reduced-order system. Moreover, the existence of computable error bounds allows an adaptive choice of the state space dimension of the approximate model. The balanced truncation methods are closely related to the proper and improper controllability and observability Gramians of system (1) that are defined by the solutions of the two dual continuous-time and two dual discrete-time projected generalized Lyapunov equations.

Note that Lyapunov-based balanced truncation, in general, does not preserve passivity in the reduced-order system. In a passivity-preserving model reduction approach, known as positive real balanced truncation, instead of the continuous-time projected Lyapunov equations we have to solve the projected generalized Riccati equations. For the DAE control system (1) that is not necessarily minimal but that has the proper transfer function $\mathbf{G}(s)$, we have the following algorithm.

Algorithm. Positive real balanced truncation for DAE systems.
Given $\mathbf{G} = [\,E,\, A,\, B,\, C,\, D\,]$, compute the reduced-order system $[\,\widetilde{E},\, \widetilde{A},\, \widetilde{B},\, \widetilde{C},\, \widetilde{D}\,]$.

1. Compute the Cholesky factors R_i and L_i of the improper controllability and observability Gramians $\mathcal{G}_{ic} = R_i R_i^T$ and $\mathcal{G}_{io} = L_i L_i^T$ by solving the projected generalized discrete-time Lyapunov equations

$$\begin{aligned} A\mathcal{G}_{ic}A^T - E\mathcal{G}_{ic}E^T &= Q_l B B^T Q_l^T, \qquad & \mathcal{G}_{ic} &= Q_r \mathcal{G}_{ic} Q_r^T, \\ A^T\mathcal{G}_{io}A - E^T\mathcal{G}_{io}E &= Q_r^T C^T C Q_r, & \mathcal{G}_{io} &= Q_l^T \mathcal{G}_{io} Q_l, \end{aligned}$$

   with $Q_r = I - P_r$ and $Q_l = I - P_l$.
2. Compute the skinny singular value decomposition $L_i^T A R_i = U\Theta V^T$, where U and V have orthonormal columns and Θ is nonsingular.
3. Compute $W_2 = L_i U \Theta^{-1/2}$, $T_2 = R_i V \Theta^{-1/2}$, $P_0 = D - CT_2W_2^T B$, $R = P_0 + P_0^T$.
4. Compute the Cholesky factors R and L of the solutions $\mathcal{X} = RR^T$ and $\mathcal{Y} = LL^T$ of the projected generalized Riccati equations

$$\begin{aligned} A\mathcal{X}E^T + E\mathcal{X}A^T + (E\mathcal{X}C^T - P_l B)R^{-1}(E\mathcal{X}C^T - P_l B)^T = 0, \qquad & \mathcal{X} = P_r \mathcal{X} P_r^T, \\ A^T\mathcal{Y}E + E^T\mathcal{Y}A + (B^T\mathcal{Y}E - CP_r)^T R^{-1}(B^T\mathcal{Y}E - CP_r) = 0, \qquad & \mathcal{Y} = P_l^T \mathcal{Y} P_l. \end{aligned}$$

5. Compute the skinny singular value decomposition

$$L^T E R = [\,U_1,\, U_2\,] \begin{bmatrix} \Pi_1 & \\ & \Pi_2 \end{bmatrix} [\,V_1,\, V_2\,]^T,$$

   where $[\,U_1,\, U_2\,]$ and $[\,V_1,\, V_2\,]$ have orthonormal columns, $\Pi_1 = \operatorname{diag}(\pi_1, \ldots, \pi_\ell)$ and $\Pi_2 = \operatorname{diag}(\pi_{\ell+1}, \ldots, \pi_r)$ with $\pi_1 \geq \ldots \geq \pi_\ell \gg \pi_{\ell+1} \geq \ldots \geq \pi_r > 0$.
6. Compute the reduced-order system

$$[\,\widetilde{E},\, \widetilde{A},\, \widetilde{B},\, \widetilde{C},\, \widetilde{D}\,] = [\,W_1^T E T_1,\, W_1^T A T_1,\, W_1^T B,\, C T_1,\, P_0\,]$$

   with $W_1 = L U_1 \Pi_1^{-1/2}$ and $T_1 = R V_1 \Pi_1^{-1/2}$.

Similarly to the standard state space case [3], one can show that the reduced-order system with the transfer function $\widetilde{\mathbf{G}}(s) = \widetilde{C}(s\widetilde{E} - \widetilde{A})^{-1}\widetilde{B} + \widetilde{D}$ is passive, and the $\mathbb{H}_\infty$-norm error bound

$$\|\widetilde{\mathbf{G}} - \mathbf{G}\|_{\mathbb{H}_\infty} \leq 2\|R^{-1}\|^2 \|\mathbf{G} + \widetilde{D}^T\|_{\mathbb{H}_\infty} \|\widetilde{\mathbf{G}} + \widetilde{D}^T\|_{\mathbb{H}_\infty} \sum_{j=\ell+1}^{r} \pi_j$$

holds where $\|\mathbf{G}\|_{\mathbb{H}_\infty} = \sup_{\omega \in \mathbb{R}} \|\mathbf{G}(i\omega)\|$ denotes the $\mathbb{H}_\infty$-norm of $\mathbf{G}$.

A major difficulty in the numerical solution of the projected Lyapunov and Riccati equations with large matrix coefficients is that the spectral projectors onto the left and right deflating subspaces corresponding to the finite and infinite eigenvalues of the pencil $\lambda E - A$ are required. However, in many applications such as control of fluid flow, electrical circuit simulation and constrained multibody systems, the matrices E and A have some special block structure. This structure can be used to construct the projectors in explicit form [1, 8].

References

[1] D. Estévez Schwarz and C. Tischendorf. *Structural analysis for electric circuits and consequences for MNA*, Int. J. Circ. Theor. Appl. **28** (2000), 131–162.

[2] K. Glover, *All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds*, Internat. J. Control **39** (1984), 1115–1193.

[3] S. Gugercin and A.C. Antoulas. *A survey of model reduction by balanced truncation and some new results*, Internat. J. Control **77** (2004), 748–766.

[4] B.C. Moore. *Principal component analysis in linear systems: controllability, observability, and model reduction*, IEEE Trans. Automat. Control **26** (1981), 17–32.

[5] T. Stykel. *Analysis and Numerical Solution of Generalized Lyapunov Equations*. Ph.D. thesis, Institut für Mathematik, Technische Universität Berlin, 2002.

[6] T. Stykel. *On criteria for asymptotic stability of differential-algebraic equations*, Z. Angew. Math. Mech. **82** (2002), 147–158.

[7] T. Stykel. *Gramian-based model reduction for descriptor systems*, Math. Control Signals Systems **16** (2004), 297–319.

[8] T. Stykel. *Low rank iterative methods for projected generalized Lyapunov equations*, Preprint 198, DFG Research Center Matheon, Technische Universität Berlin, 2004.

DAEs and PDEs for the Simulation of Shape Memory Behavior

Gunnar Teichelmann

(joint work with Bernd Simeon)

Shape Memory Alloy (SMA) materials have an enormous potential in technological applications like aviation or medicine among others. This talk aims at their use as temperature controlled actuators in mechatronic applications. All this calls for simulation tools and techniques that are able to describe the relevant effects of SMA behavior. Both mathematical models and appropriate numerical simulation schemes have to be developed. In our case the model described by Helm [1] presents major challenges since it consists of a heterogeneous coupled system of partial differential and differential-algebraic equations (PDAEs) where continuum models describe the evolution of deformation and temperature. State of the art solution methods in most cases use return mapping algorithms comparable to low order implicit integration schemes with fixed stepsize. So far, the use of finite element methods to simulate shape memory behavior has mainly focused on the isothermal case. Only few papers about the thermomechanic coupling exist. By the interpretation of the multiphysical problem as PDAE the range of applicable solution methods widens. After semidiscretization in space the resulting DAE is open to be treated with appropriate time integration schemes with step size control.

A material point x in referential coordinates is mapped to its position on the deformed domain by the deformation function $\vartheta(t,x) = x + u(t,x)$ with the displacement field u, see [2]. We formulate the quasistationary momentum balance (1) with the symmetric stress tensor σ and the density of body forces β. Furthermore, we have mixed Dirichlet and Neumann boundary conditions and of course consistent initial conditions. The kinematic relation between displacement u and strain ϵ is represented by the linearized Lagrangian strain, while the total strain

is decomposed by $\epsilon = \epsilon_p + \epsilon_e$ into plastic and elastic strain. The relation between stress σ and elastic strain ϵ_e is then given by a generalized Hooke's law, depending also on the temperature. There are 6 ways for the volume fractions of temperature induced martensite z_{TIM}, stress induced martensite and austenite to change from one to another. The decision which transition is active depends mainly on the temperature and its rate, the martensite fractions, the loading conditions, the internal stress X and the value of the yield function. The variables z_{TIM}, X and ϵ_p are internal values, that are comprised in the vector α. They are described by evolution equations (3), whose right hand sides discontinuously depend on the actual phase transition. Also the heat equation (2) is part of the mathematical model where the source term $f(t, x, \alpha, \dot{\alpha}, \eta)$ depends on all the variables and their change rates in time. Until now the system reads

$$0 = \operatorname{div} \sigma(u, \epsilon_p) + \beta(t), \tag{1}$$

$$\rho c_0 \, \dot{\theta} = \lambda \Delta \theta + f(t, x, \alpha, \dot{\alpha}, \eta), \tag{2}$$

$$\dot{\alpha} = \Gamma(\alpha, \sigma, \theta). \tag{3}$$

The balance of linear momentum depends both on the internal variable ϵ_p and the temperature θ while the heat equation needs information about the strains ϵ and the strain rates $\dot{\epsilon}$. In contrast to $\dot{\epsilon}_{\mathrm{p}}$ the strain rate $\dot{\epsilon}$ is not available in a quasistationary computation of the displacements. To overcome this problem we introduce a new DAE with a Lagrangian multiplier η in the following sense

$$\dot{\epsilon} = \eta, \quad 0 = \epsilon - \frac{1}{2}(\nabla u + \nabla u^T).$$

Coupling this DAE to our existing differential equations system, the quantity $\dot{\epsilon} = \eta$ is now available on the right hand side. Note that this additional DAE is of differentiation and perturbation index 2 and therefore the choice of an appropriate integration scheme is crucial. At integration the tolerance values for the algebraic part of this DAE can or should be chosen relatively coarse, depending on the integration scheme. Detailed information can additionally be found in [3, 4].

For solution purpose we first use a finite element approach for semidiscretization in space of the quasistationary momentum balance and the heat equation. In this course the discretization points of the internal variables in space and of the additional index 2 DAE show to be the gauss nodes of the fem grid. This procedure leads to the semidiscretized system

$$\begin{aligned}
0 &= K \cdot q - b - Q \cdot \bar{\epsilon}_p, \\
M_\theta \cdot \dot{q}_\theta &= -K_\theta(q) \cdot q_\theta + b_\theta(\alpha, \dot{\alpha}, \eta), \\
\dot{\alpha}_i &= \Gamma(\alpha_i, \sigma(q, \xi_i)), \\
\dot{\epsilon}_i &= \eta_i, \qquad\qquad i = 1 \ldots k, \\
0 &= \epsilon_i - B(\xi_i) \cdot q
\end{aligned}$$

of coupled ODE and DAE.

The discontinuous right hand side of the evolution prevents implicit integration schemes to succeed. The smoothing of these discontinuities allows implicit integration with the drawback of small stepsizes and an increased stiffness in these regions. In a benchmark simulation of a onedimensional shape memory wire the BDF2 scheme performed well. With 570 successful steps the number of 186 convergence failures seems to be high, but they are due to the phase transitions. At these points the evolution equations for the internal variables are changed and thus the step size in time is reduced considerably. The error of the strain rate that is determined by the index 2 DAE stays inside the integration tolerance.

References

[1] D. Helm, *Formgedächtnislegierungen: Experimentelle Untersuchung, phänomenologische Modellierung und numerische Simulation der thermomechanischen Materialeigenschaften*, PhD thesis at Universität Gesamthochschule Kassel, 2001

[2] J.E. Marsden and T.J.R. Hughes, *Mathematical Foundations of Elasticity*, Dover Publications, 1983

[3] G. Teichelmann and B. Simeon, *Numerical Simulation of SMA Actuators*, Progress in Industrial Mathematics at ECMI 2004, Springer Verlag (2006), 647–651

[4] G. Teichelmann, B. Simeon and D. Helm, *Modeling and Simulation of Shape Memory Behavior*, ARGESIM Report **30** (2006)

Abstract Differential-Algebraic Equations

Caren Tischendorf

The simulation of complex systems describing different physical effects becomes more and more of interest in various applications, for instance, in chip design, in the development of micro-electro-mechanical systems (MEMS), in structural mechanics, in biomechanics and in medicine. The modeling of complex processes often lead to coupled systems that are composed of ordinary differential equations (ODEs), differential-algebraic equations (DAEs) and partial differential equations (PDEs).

Such coupled systems can be regarded in the general framework of abstract differential-algebraic equations of the form

$$\mathcal{A}(u,t)\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{D}(u,t) + \mathcal{B}(u,t) = 0, \qquad t \in [t_0, T].$$

This equation is to be understood as an operator equation with operators $\mathcal{A}(\cdot,t)$, $\mathcal{D}(\cdot,t)$ and $\mathcal{B}(\cdot,t)$ acting in real Hilbert spaces where $u : [t_0,T] \to X$ is the solution belonging to a problem adapted space.

If the Hilbert spaces are chosen to be the finite dimensional space $\mathbb{R}^m$, then we obtain a differential-algebraic equation. Choosing $\mathcal{A}$ and $\mathcal{D}$ as the natural embedding operators, we obtain an evolution equation. If, additionally, $\mathcal{B}$ is a second-degree differential operator in space, it leads to a parabolic differential equation. For elliptic differential equations, the operators $\mathcal{A}$ and $\mathcal{D}$ are identically zero.

For most coupled systems, the operators $\mathcal{A}$ and $\mathcal{D}$ are neither identically zero nor invertible on the time interval $[t_0, T]$. For example, coupled circuit and device models in chip design [14] have leading term operators of the form

$$\mathcal{A} = \begin{pmatrix} A & 0 & 0 \\ 0 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & \mathcal{I} \end{pmatrix}, \qquad \mathcal{D}(u,t) = \begin{pmatrix} q(u_1,t) \\ \mathcal{R}u_2 \\ u_3 \end{pmatrix}$$

where A is a constant, finite-dimensional, singular matrix, I is the identity, $\mathcal{I}$ is the natural embedding operator, $q : \mathbb{R}^n \times [t_0, T] \to \mathbb{R}^m$ and $\mathcal{R}$ is a boundary integral operator.

A general theory of abstract differential-algebraic equations (ADAEs) does not exist and can not be expected to be given considering alone the complexity of problems simulating partial differential equations. However, special classes of ADAEs have recently been successfully analyzed and simulated, see e.g. [1]-[14]. We presented a short overview of the treated classes and discussed basic ideas of the different approaches to handle coupled problems.

In particular, we considered solvability and perturbation results for linear ADAEs with time-constant coefficients using Laplace transformation and decoupling techniques [11, 12], for linear ADAEs with monotone, time dependent coefficients by a Galerkin approach [14], for linear elliptic and parabolic PDAEs using Gårding-type inequalities [10], for nonlinear ADAEs from coupled circuit and device simulation by fixed point arguments and decoupling techniques [1]-[4] as well as for flexible multibody systems employing saddle point arguments [13].

References

[1] G. Alïi A. Bartel, M. Günther and C. Tischendorf, *Elliptic partial differential-algebraic multiphysics models in electrical network design*, M^{3}AS, **13:9** (2003), 1261–1278.

[2] G. Alïi A. Bartel and M. Günther, *Parabolic differential-algebraic models in electrical network design*, SIAM J. Mult. Model. Sim., **4:3** (2005), 813–838.

[3] M. Bodestedt: *Index of coupled systems in circuit simulation.*, Licentiate dissertation, Lund University, Sweden (2004).

[4] M. Bodestedt and C. Tischendorf, *PDAE models of integrated circuits and perturbation analysis*, Preprint 2004–8, Institute of Math., Humboldt Univ. of Berlin, Germany, (2004). To appear in *Math. Comput. Model. Dyn. Syst.*

[5] S.L. Campbell and W. Marszalek *The index of infinite dimensional implicit systems*, Math. Comput. Model. Dyn. Syst., **5:1** (1999), 18-42.

[6] C. Eichler-Liebenow *Numerical treatment of higher space dimensional parabolic differential equations by linearly implicit splitting methods and of linear partial algebraic equations*, PhD thesis, Martin-Luther-Univ. Halle-Wittenberg, Germany (1999).

[7] A. Favini and A. Yagi *Degenerate differential equations in Banach spaces*, Pure and Applied Mathematics, Marcel Dekker, New York (1999).

[8] R. Lamour, R. März and C. Tischendorf *PDAEs and further mixed systems as abstract differential-algebraic systems*, Preprint 01–11, Institute of Math., Humboldt Univ. of Berlin, Germany (2001).

[9] W.S. Martinson and P.I. Barton *Index and characteristic analysis of linear PDAE systems*, SIAM J. Sci. Comput. **24:3** (2002), 905–923.

[10] J. Rang: Stability estimates and numerical methods for degenerate parabolic differential equations, PhD thesis, TU Clausthal, Germany (2004).
[11] T. Reis: Systems theoretic aspects of PDAEs and applications to electrical circuits, PhD thesis, TU Kaiserslautern, Germany (2006).
[12] T. Reis and C. Tischendorf *Frequency domain methods and decoupling of linear infinite dimensional differential algebraic systems*, J. Evol. Equ. **5:3** (2005), 357–385.
[13] B. Simeon *Numerische Simulation gekoppelter Systeme von partiellen und differential-algebraischen Gleichungen in der Mehrkörperdynamik*, Fortschritt-Berichte VDI, Reihe 20, Nr. 325, Düsseldorf, VDI-Verlag (2000).
[14] C. Tischendorf *Coupled systems of differential-algebraic and partial differential equations in circuit and device simulation. Modeling and numerical analysis*, Habilitation thesis, Humboldt Univ. of Berlin, Germany (2004).

The IMEX Runge-Kutta-Chebyshev Method for Stiff Advection-Diffusion-Reaction Problems

J.G. Verwer

This lecture is devoted to the time integration of stiff, nonlinear advection-diffusion-reaction PDE problems. Adopting the method of lines approach we assume that the PDE system with its boundary conditions has been spatially discretized, and thus we focus on ODE systems

$$w'(t) = F(t, w(t)), \quad t > 0, \qquad w(0) = w_0, \tag{1}$$

representing semi-discrete advection-diffusion-reaction problems. In most practical applications the dimension of this ODE system is huge, especially for multi-space dimensional PDEs and/or PDE systems with many reacting species. The huge dimension and the simultaneous occurrence of advection, diffusion and reaction terms and stiffness can severely complicate the use of standard implicit integrators leaning on modified Newton and (preconditioned iterative) linear solvers. On the other hand, the stiffness induced by diffusion and reaction terms rules out easy-to-use standard explicit solvers. This delineates our research question: how to realize easy-to-use, robust and efficient time stepping for this sort of semi-discrete PDEs.

Decoupling the three processes from one another generally simplifies matters. Most simple is to use operator (time) splitting by which advection, diffusion and reactions can be sequentially and independently solved with integrators tuned for the three different parts, see Ch. IV of [2]. A drawback is that operator splitting can give rise to large splitting errors for operators exhibiting slow and fast time scales that nearly balance. In particular, operator splitting is not exact for steady states which is a disadvantage for transient problems running into steady state. In this respect, decoupling through the implicit-explicit (IMEX) approach is more subtle and preserves transient balances.

In [5] we have proposed a Runge-Kutta-Chebyshev (RKC) method of the IMEX type treating modestly stiff diffusion terms explicitly and highly stiff reaction terms giving rise to real eigenvalues implicitly. The Fortran90 code IRKC implements the IMEX method [3]. The explicit part of this IMEX method closely resembles the first RKC method due to van der Houwen & Sommeijer [1]. This method

is an explicit, second-order stabilized Runge-Kutta method with a real stability boundary equal to $\approx 0.66s^2$, s being the number of stages. A strong property is that s can be taken arbitrarily large without internal error growth. The implicit part of the IMEX method has been designed such that for the stiff reaction terms the method is unconditionally stable (linear stability as used in the stiff ODE field). Furthermore, if the reaction terms do not imply spatial dependence, in the method they remain uncoupled over space grids enabling a fast computation of the stiff terms (a single ODE system per grid box of a dimension equal to the number of PDEs).

In [6] we have further extended our explicit method with the aim to also include advection terms. Herewith our final goal is an efficient implicit-explicit RKC integration of advection-diffusion-reaction PDE problems in a manner that advection and diffusion terms are treated simultaneously and explicitly and the highly stiff reaction terms implicitly.

This talk reviews the developments towards this goal accompanied with several numerical illustrations.

References

[1] P.J. van der Houwen, B.P. Sommeijer (1980), *On the internal stability of explicit, m-stage Runge-Kutta methods for large m-values*. Z. Angew. Math. Mech. 60, pp. 479–485.

[2] W. Hundsdorfer, J.G. Verwer (2003), *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*, Springer Series in Computational Mathematics, Vol. 33, Springer, Berlin.

[3] L.F. Shampine, B.P. Sommeijer, J.G. Verwer (2006), *IRKC: An IMEX solver for stiff diffusion-reaction PDEs*, J. Comput. Appl. Math., to appear.

[4] B.P. Sommeijer, L.F. Shampine, J.G. Verwer (1997), *RKC: An explicit solver for parabolic PDEs*. J. Comput. Appl. Math. 88, pp. 315–326.

[5] J.G. Verwer, B.P. Sommeijer (2003), *An implicit-explicit Runge-Kutta-Chebyshev scheme for diffusion-reaction equations*. SIAM J. Sci. Comput. 25, pp. 1824-1835.

[6] J.G. Verwer, B.P. Sommeijer, W. Hundsdorfer (2004), *RKC time-stepping for convection-diffusion-reaction problems*. J. Comput. Phys. 201, pp. 61-79.

General Linear Methods for Integrated Circuit Design

Steffen Voigtmann

Today electronic devices play an important part in everybody's life. In particular, there is an ongoing trend towards using mobile devices such as cell phones, laptops or PDAs. Integrated circuits for these kind of applications are mainly produced in CMOS technology (complementary metal-oxide semiconductor). CMOS circuits use almost no power when they are not active and thus, combining negatively and positively charged transistors, they draw power only when switching polarity.

Circuit simulation is one of the key technologies enabling a further increase in performance and memory density. One important analysis type in circuit simulation is the transient analysis of layouts on varying input signals. Based on schematics or netlist descriptions of electrical circuits the corresponding model equations are automatically generated using the modified nodal analysis (MNA).

This network approach preserves the topological structure of the circuit but does not lead to a minimal set of unknowns. Hence the resulting model consists of differential algebraic equations (DAEs)

$$A\,\dot{q}\big(x(t),t\big) + b\big(x(t),t\big) = 0. \tag{1}$$

The vector $x(t) \in \mathbb{R}^m$ comprises all node potentials and some branch currents while $q(x,t) \in \mathbb{R}^n$ represents charges and fluxes [2, 7]. Note that (1) has a properly stated leading term in the sense of [4].

Typically MNA equations suffer from poor smoothness properties due to the model equations of modern transistors but also due to e.g. piecewise linear input functions. Similarly, time constants of several orders of magnitudes give rise to stiff equations and low order A-stable methods need to be used.

The further miniaturisation of electrical devices drives simulation methods for circuit DAEs to their limits. Due to the reduced signal/noise ratio, stability questions become more and more important for modern circuits. Thus there is a strong need to improve stability properties of existing methods such as the combination of BDF and trapezoidal rule. There are fully implicit Runge-Kutta methods that exhibit much better stability properties. However, these methods are currently not attractive for industrial circuit simulators due to their high computational costs.

In order to cope with these difficulties, general linear methods (GLMs) are studied for integrated circuit design. These methods were introduced by John Butcher to provide a framework covering, among others, both linear multistep and Runge-Kutta methods. They enable the construction of new methods with improved convergence and stability properties [1].

A general linear method is characterised by four matrices $\mathcal{M} = [\mathcal{A},\mathcal{U},\mathcal{B},\mathcal{V}]$ satisfying $\mathcal{A} \in \mathbb{R}^{s\times s}$, $\mathcal{U} \in \mathbb{R}^{s\times r}$, $\mathcal{B} \in \mathbb{R}^{r\times s}$, $\mathcal{V} \in \mathbb{R}^{r\times r}$. The integer s is referred to as the number of internal stages, while r denotes the number of external stages. In order to proceed from the timepoint t_n to $t_{n+1} = t_n + h$ using a stepsize h, r input quantities $q_j^{[n]} \in \mathbb{R}^n$, $j = 1,\ldots,r$ are used to compute s stage approximations $X_i \approx x(t_n + c_i h) \in \mathbb{R}^m$, $i = 1,\ldots,s$, at intermediate timepoints. An updated vector $q^{[n+1]}$ is passed on to the next step. The interrelation of the various quantities is given by the numerical scheme

$$A\,Q_i' + b(X_i, t_n + c_i h) = 0, \qquad \begin{aligned} Q &= h\,(\mathcal{A}\otimes I_n)\,Q' + (\mathcal{U}\otimes I_n)\,q^{[n]} \\ q^{[n+1]} &= h\,(\mathcal{B}\otimes I_n)\,Q' + (\mathcal{V}\otimes I_n)\,q^{[n]} \end{aligned} \tag{2}$$

where $i = 1,\ldots,s$ and

$$Q = \begin{bmatrix} q(X_1, t_n + c_1 h) \\ \vdots \\ q(X_s, t_n + c_s h) \end{bmatrix}, \quad Q' = \begin{bmatrix} Q_1' \\ \vdots \\ Q_s' \end{bmatrix}, \quad q^{[n]} = \begin{bmatrix} q_1^{[n]} \\ \vdots \\ q_r^{[n]} \end{bmatrix}, \quad q^{[n+1]} = \begin{bmatrix} q_1^{[n+1]} \\ \vdots \\ q_r^{[n+1]} \end{bmatrix}.$$

The computational complexity of this scheme is mainly determined by the structure of the matrix $\mathcal{A}$. If $\mathcal{A}$ has a diagonally implicit structure, the stages X_i can be evaluated sequentially such that the computational costs are significantly reduced. One of the key advantages of general linear methods over Runge-Kutta schemes

is the fact that diagonally implicit methods with high stage order are possible [9]. Hence, in spite of the diagonally implicit structure, there will be no order reduction for the index-2 components.

Studying general linear methods for DAEs requires a thorough analysis of these equations. Using the framework of the tractability index [4] it is possible to derive a decoupling procedure for nonlinear index-2 DAEs. In contrast to similar decouplings for linear equations, the inherent dynamics is characterised by an implicit index-1 equation.

Theorem 4. *Let* (1) *be a regular index-2 DAE with a properly stated leading term. Assume that* $N_0 \cap S_0$ *does not depend on* x*. Then* (1) *is locally equivalent to index equation*

$$u' = f(u, z', t), \qquad v = g(u, t), \tag{3a}$$

$$z = \mathbb{z}(u, t), \qquad w = \mathbb{w}(u, v', t), \tag{3b}$$

where (3a) *represents an implicit index-1 system. The solution of* (1) *is given by* $x = D^- u + z + w$ *where* D^- *is a generalised reflexive inverse of* $\frac{\partial q}{\partial x}$*.*

Details on the subspaces N_0, S_0, the involved functions and the required smoothness assumptions are given in [7]. The proof given there presents a decoupling procedure that transforms (1) into (3). It is stressed that this approach does not require the derivative array and only mild smoothness assumptions are made. In case of linear DAEs, (3a) reduces to the inherent regular ODE derived in [4].

This new decoupling procedure for nonlinear index-2 DAEs allows to prove existence and uniqueness results requiring only mild smoothness properties. Additionally, general linear methods for index-2 DAEs can be studied by investigating implicit index-1 equations first.

Order conditions ensuring order p behaviour for the local discretisation error can be derived using rooted trees. The approach is similar to the one taken for Runge-Kutta methods in [3]. Notice, however, that these results on Runge-Kutta schemes present only a subset of the order conditions for general linear methods. Due to the multivalue nature of the methods, additional order conditions have to guarantee the required order for all components of $q^{[n]}$. For methods in Nordsieck form, where the input quantities $q_{j+1}^{[n]} \approx h^j \frac{d^j}{dt^j} q\big(x(t), t\big)$, $j = 0, \ldots, r-1$, approximate scaled derivatives of the exact solution, the full set of order conditions has been derived in [7].

The order conditions and convergence results can be transferred to the general index-2 equation (1) using the decoupling procedure discussed above.

Theorem 5. *Let* $\mathcal{M} = [\mathcal{A}, \mathcal{U}, \mathcal{B}, \mathcal{V}]$ *be a GLM in Nordsieck form. Assume that*

- $\mathcal{M}$ *has order* p *for implicit index-1 DAEs* (3a)*,*
- $\mathcal{V}$ *is power bounded and* $M_\infty = \mathcal{V} - \mathcal{B}\mathcal{A}^{-1}\mathcal{U}$ *nilpotent with* $M_\infty^k = 0$*,*
- $\mathcal{M}$ *is stiffly accurate, i.e.* $e_s^\top \mathcal{A} = e_1^\top \mathcal{A}$ *and* $e_s^\top \mathcal{U} = e_1^\top \mathcal{V}$*,*
- $\mathcal{M}$ *has stage order* q *for ordinary differential equations.*

Then, after k *steps,* $\mathcal{M}$ *is convergent with order* $\min(p, q)$ *for regular index-2 DAEs* $A\,\dot{q}\big(x(t), t\big) + b\big(x(t), t\big) = 0$ *with a properly stated leading term.*

A proof and further details can be found in [7].

The full set of requirements on methods for nonlinear DAEs has to be taken into account when constructing practical methods. A test implementation GLIMDA [8] has been developed that employs General LInear Methods for Differential Algebraic equations. The code implements a variable-stepsize, variable-order approach, where methods of order 1,2 and 3 are used.

The preliminary code GLIMDA based on general linear methods seems to be competitive with BDF and Runge-Kutta solvers. By construction GLIMDA has advantages for MNA equations. Hence there is strong evidence that general linear methods can be used efficiently for solving differential algebraic equations in integrated circuit design.

References

[1] J.C. Butcher, *Numerical methods for ordinary differential equations*, John Wiley & Sons Ltd., Chichester (2003)

[2] U. Feldmann, M. Günther, J. ter Maten, *Modelling and Discretization of Circuit Problems*, Handbook of Numerical Analysis, Volume XIII: Numerical Methods in Electromagnetics, North-Holland (2005)

[3] A. Kværnø, *Runge-Kutta methods applied to fully implicit differential-algebraic equations of index* 1, Math. Comp. **54** no. 190 (1990), 583–625

[4] R. März, *Differential algebraic systems with properly stated leading term and MNA equations*, Modeling, simulation, and optimization of integrated circuits (Oberwolfach, 2001), Internat. Ser. Numer. Math. **146**, Birkhäuser Basel (2003), 135–151

[5] S. Voigtmann, *General linear methods for nonlinear DAEs in circuit simulation*, Scientific Computing in Electrical Engineering (2004), to appear.

[6] S. Voigtmann, *Accessible criteria for the local existence and uniqueness of DAE solutions*, Technical Report 214, MATHEON (2005)

[7] S. Voigtmann, *General Linear Methods for Integrated Circuit Design*, PhD thesis, Humboldt Universität zu Berlin (2006), in preparation

[8] S. Voigtmann, GLIMDA – *General LInear Methods for Differential Algebraic equations*, `http://www.math.hu-berlin.de/~steffen/software.html`

[9] W. Wright, *General linear methods with inherent Runge-Kutta stability*, PhD thesis, The University of Auckland, New Zealand (2003)

Collocation Methods for Index-1 DAEs with a critical point

EWA B. WEINMÜLLER

(joint work with O. Koch, R. März, D. Praetorius)

Model problem. We investigate the convergence behavior of collocation schemes applied to approximate solutions of index-1 DAEs, including the case when a critical point of $1-A$ type is present, see [6] and [5] for more technical details. The underlying analytical problem is the linear system of DAEs,

$$A(t)(D(t)x(t))' + B(t)x(t) = g(t), \quad t \in (0,1], \tag{1}$$

where $A(t) \in \mathbb{R}^{m\times n}$, $D(t) \in \mathbb{R}^{n\times m}$, $B(t) \in \mathbb{R}^{m\times m}$ and $g(t)$, $x(t) \in \mathbb{R}^m$ with $n \le m$. We assume that $D(t) \equiv D$ is a constant matrix and that the matrices A, B and the inhomogeneity g are at least continuous, A, B, $g \in C[0,1]$.

Example: The following two dimensional problem belongs to class (1) and has a solution $x_1(t) = -\frac{6t+1}{2}e^{5t}$, $x_2(t) = -\frac{8t+1}{2}e^{5t}$:

$$\begin{pmatrix}1\\1\end{pmatrix}(1,-1)\begin{pmatrix}x_1(t)\\x_2(t)\end{pmatrix}'+\begin{pmatrix}2 & 0\\0 & t+2\end{pmatrix}x(t) \;=\; \begin{pmatrix}-te^{5t}\\-\frac{8t+7}{2}te^{5t}\end{pmatrix}. \tag{2}$$

We study systems (1) with properly stated leading term, cf. [1]. This means that A and D are well matched, i.e., $\ker A(t) \oplus \operatorname{im} D(t) = \mathbb{R}^n$, $t \in (0,1]$, and there exists a projector function $R \in C^1(0,1]$ which realizes this splitting. Here, we assume that $\ker(A(t)) = \{0\}$, $t \in (0,1]$ and $\operatorname{im}(D) = \mathbb{R}^n$. Let Q_0 be a projector onto $N_0 := \ker(A(t)D) \equiv \ker(D)$ and let us define $P_0 := I - Q_0$. In our case, since the matrix D is constant, $R = I$ for $t \in (0,1]$, and Q_0, P_0 are constant, we regard all projections as extended to the interval $[0,1]$. In order to describe the boundary/initial conditions which are necessary and sufficient for (1) to be well-posed, we decouple the system using techniques from [1]. To this end, we introduce the matrices $G_0(t) := A(t)D$, $G_1(t) := G_0(t) + B(t)Q_0$ and allow a critical point at $t = 0$, where G_1 may become singular, i.e. $G_1(t)$ is non-singular on $(0,1]$. The decoupled system reads:

$$u'(t) + DG_1^{-1}(t)B(t)D^-u(t) = DG_1^{-1}(t)g(t), \quad t \in (0,1], \tag{3}$$

$$Q_0x(t) = -Q_0G_1^{-1}(t)B(t)D^-u(t) + Q_0G_1^{-1}(t)g(t), \quad t \in (0,1], \tag{4}$$

where $u(t) := Dx(t)$ are differential and $Q_0x(t)$ are algebraic components of the solution $x(t)$, and D^- is a reflexive generalized inverse of D. We now rewrite (3) and obtain a system of singular ODEs with a singularity of the first kind[1],

$$u'(t) - \frac{1}{t}M(t)u(t) = f(t), \;\; t \in (0,1], \tag{5}$$

where $M(t)/t := -DG_1^{-1}(t)B(t)D^-$, $f(t) := DG_1^{-1}(t)g(t)$. Let us assume that $M \in C^1[0,1]$ and $f \in C[0,1]$. Then we can use the theory given in [3] to augment (5) by a set of initial[2] conditions necessary and sufficient for $u \in C[0,1]$. In case that $M(0)$ has zero eigenvalues or eigenvalues with negative real parts, u needs to satisfy $u(0) = \gamma$, where $\gamma \in \ker M(0)$. Finally, if the right-hand side in (4) is continuous on $[0,1]$, then there exists a unique, continuous solution of the following IVP:

$$A(t)Dx'(t) + B(t)x(t) = g(t), \quad t \in (0,1], \tag{6}$$

$$Dx(0) = \gamma, \;\; Q_0x(0) = \lim_{t\to 0}(-Q_0G_1^{-1}(t)B(t)D^-\gamma + Q_0G_1^{-1}(t)g(t)) =: Q_0x_0. \tag{7}$$

Collocation scheme. We now turn to the numerical treatment of the IVP (6), (7). We first introduce a mesh $\Delta := (\tau_0, \tau_1, \ldots, \tau_N)$, with $h_i := \tau_{i+1} - \tau_i$, $i = 0, \ldots, N-1$, $\tau_0 = 0$, $\tau_N = 1$, such that $h_i \le h$. In each subinterval $J_i = [\tau_i, \tau_{i+1}]$, we place m distinct collocation points, $\tau_i < t_{i,j} < \tau_{i+1}$, $j = 1, \ldots, m$. We approximate $x(t)$ by a function $p(t) = p_i(t)$, $t \in J_i$, where $p \in \mathbf{B}_m$, and $\mathbf{B}_m$ is

[1]Singularity of the first kind arises when we assume that $t = 0$ is an algebraically simple zero of the determinant of $G_1(t)$.

[2]We restrict our attention to IVPs in this talk.

the Banach space of globally continuous, piecewise polynomial functions of degree $\leq m$ equipped with the maximum norm. The defining equations for p are, $j = 1, \ldots, m,\ i = 0, \ldots, N-1$,

$$A(t_{i,j})Dp'(t_{i,j}) + B(t_{i,j})p(t_{i,j}) = g(t_{i,j}), \tag{8}$$

$$Dp(0) = \gamma, \quad Q_0 p(0) = Q_0 x_0. \tag{9}$$

Note, that the numerical method is applied to the IVP (6), (7) in its original form. We first show that $p \in \mathbf{B}_m$ exists and is unique. Decoupling (8) yields a collocation scheme for the differential components of p, $q(t) := Dp(t)$, and it follows from [4] that $q(t) \in \mathbf{B}_m$ exists and is unique. Then, it is easy to see that $Q_0 p(t) \in \mathbf{B}_m$ exists and is unique and consequently, this also holds for $p(t) \in \mathbf{B}_m$.

In order to derive the error bounds for the solution p, we introduce an error function $e \in \mathbf{B}_m$ defined by, $j = 1, \ldots, m,\ i = 0, \ldots, N-1$,

$$e'(t_{i,j}) = x'(t_{i,j}) - p'(t_{i,j}), \quad e(0) = 0. \tag{10}$$

Standard results for interpolation, see [2], yield the estimate for the interpolation error $e'(t) = x'(t) - p'(t) + P_0 O(h^k) + Q_0 O(h^l)$. Integrating this expression, we obtain $e(t) = x(t) - p(t) + t(P_0 O(h^k) + Q_0 O(h^l))$ provided that $P_0 x \in C^{\tilde{k}+1}[0,1]$ or equivalently $Dx \in C^{\tilde{k}+1}[0,1]$ and $Q_0 x \in C^{\tilde{l}+1}[0,1]$, where $k := \min\{\tilde{k}, m\}$ and $l := \min\{\tilde{l}, m\}$. Now, the error e satisfies the collocation scheme

$$A(t_{i,j})De'(t_{i,j}) + B(t_{i,j})e(t_{i,j}) = t_{i,j}B(t_{i,j})(P_0 O(h^k) + Q_0 O(h^l)), \quad e(0) = 0$$

which we again decouple. According to [4] we have $e_{\text{diff}} := De(x) = tO(h^k)$, and we can use this information to estimate $Q_0 e(t)$. Finally, $x(t) - p(t) = O(h^{\min\{l,k\}})$ follows. For details, the reader is referred to [5].

Numerical experiment. Finally, we present some numerical results to illustrate the theory.

Example: For the test problem specified in (2) the algebraic components and the differential components are given by $Q_0 x(t) = (x_2(t), x_2(t))^T$ and $P_0 x(t) = (x_1(t) - x_2(t), 0)^T = (Dx(t), 0)^T$, respectively. Moreover,

$$G_0(t) = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}, \ G_1(t) = \begin{pmatrix} 1 & 1 \\ 1 & t+1 \end{pmatrix}, \ G_1^{-1}(t) = \frac{1}{t}\begin{pmatrix} t+1 & -1 \\ -1 & 1 \end{pmatrix}.$$

The inherent singular IVP has the form $u'(t) - (-4-2t)/t\, u(t) = (7t+5)e^{5t},\ u(0) = 0$ and $u(t) = te^{5t}$. Since $u(t), x_2(t) \in C^\infty[0,1]$, we have $\tilde{k} = \tilde{l} = \infty$, and thus we expect to see the order of convergence m being the stage order of the method. In the table below we display the estimated convergence order for $m = 2$ equidistantly spaced collocation points, left column, and $m = 2$ Gaussian points, right column. The maximum norm of the global error has been calculated at the meshpoints τ_i.

Mesh	Error for x, equidistant coll.			Error for x, Gaussian coll.		
N	error	order	const.	error	order	const.
10	$1.322e+01$			$9.932e+00$		
20	$3.345e+00$	2.0	$1.271e+03$	$2.511e+00$	2.0	$9.572e+02$
40	$8.409e-01$	2.0	$1.307e+03$	$6.310e-01$	2.0	$9.819e+02$
80	$2.110e-01$	2.0	$1.320e+03$	$1.583e-01$	2.0	$9.906e+02$
160	$5.291e-02$	2.0	$1.324e+03$	$3.970e-02$	2.0	$9.936e+02$
320	$1.327e-02$	2.0	$1.326e+03$	$9.954e-03$	2.0	$9.948e+02$

Mesh	Error for u, equidistant coll.			Error for u, Gaussian coll.		
N	error	order	const.	error	order	const.
10	$7.122e-01$			$3.172e-02$		
20	$1.729e-01$	2.0	$7.847e+01$	$2.029e-03$	4.0	$2.937e+02$
40	$4.290e-02$	2.0	$7.152e+01$	$1.275e-04$	4.0	$3.165e+02$
80	$1.070e-02$	2.0	$6.935e+01$	$7.984e-06$	4.0	$3.240e+02$
160	$2.675e-03$	2.0	$6.871e+01$	$4.992e-07$	4.0	$3.262e+02$
320	$6.685e-04$	2.0	$6.853e+01$	$3.120e-08$	4.0	$3.269e+02$

The numerical results are in good agreement with the theory. The superconvergence does not hold in general although it can be observed for the differential components here. However, if we rerun the test for $m = 3$ Gaussian points, we see the $O(h^4)$ convergence for u again, and not the superconvergence behavior $O(h^6)$, see [5].

Conclusion. The concept of a properly stated leading term and the associated decoupling technique are powerful tools which we were able to utilize in the convergence proof of a collocation method applied to approximate solutions of singular DAEs. The results presented here will be subject to generalizations, such as variable matrix D, general spectrum of $M(0)$, nonlinear homogeneity, and more involved types of critical points.

References

[1] K. Balla, R. März *A unified approach to linear differential equations and their adjoint equations*, J. Anal. Appl. **21** (2003), 175–200.

[2] F. B. Hildebrand, *Introduction to Numerical Analysis*, McGraw-Hill, New York, 2nd edition, 1974.

[3] F. de Hoog, R. Weiss, *On the boundary value problems for systems of ordinary differential equations with a singularity of the first kind*, SIAM J. Math. Anal. **11** (1980), 41–60.

[4] F. de Hoog, R. Weiss, *Collocaton methods for singular boundary value problems*, SIAM J. Numer. Anal. **15** (1978), 198–217.

[5] O. Koch, R. März, D. Praetorius, E. B. Weinmüller *Collocation methods for Index-1 DAEs with a Singularity of the First Kind*, in preparation.

[6] R. März, R. Riaza *On linear algebraic-differential equations with properly stated leading term. II: Critical points*, Preprint Humboldt Unversity Berlin **23-04** (2004).

Stochastic DAEs in circuit simulation

Renate Winkler

One of the challenges of the downscaling in the production of electronic chips is the small signal-to-noise-ratio. In several applications the noise influences the system behaviour in an essentially nonlinear way such that linear noise analysis is no longer satisfactory and transient noise analysis, i.e., the integration of noisy

systems in the time domain, becomes necessary.
We deal with the thermal noise of resistors as well as the shot noise of semiconductors. Both are modelled by additional sources of additive or multiplicative Gaussian white noise currents. The thermal noise current of an resistance R at temperature T is given by Nyquist's formula $I_{th} = \sqrt{\frac{2kT}{R}}\xi(t)$. Here $\xi(t)$ denotes Gaussian white noise, and $k = 1.3806 \times 10^{-23}$ is Boltzmann's constant. The shot noise current through an pn-junctions with deterministic current I_{det} is given by Schottky's formula $I_{sh} = \sqrt{q_e I_{det}}\ \xi(t)$, see, e.g., [4]. Here $q_e = 1.602 \times 10^{-19}$ denotes the elementary charge. In both cases the noise intensities contain a small parameter.

Combining Kirchhoff's Current law with the element characteristics and using the charge-oriented formulation yields a stochastic differential algebraic equation (SDAE) of the form

$$A\frac{d}{dt}x(t) + f(t, x(t)) + \sum_{r=1}^{m} g_r(t, x(t))\xi_r(t) = 0 \tag{1}$$

where A is a constant singular matrix determined by the topology of the electrical network and ξ is an m-dimensional vector of independent Gaussian white noise sources. See, e.g., [6, 7] for the deterministic case and [5, 16] for the stochastic case. One has to deal with a large number of equations as well as of noise sources. Compared to the other quantities the noise intensities $g_r(t, x)$ are small.
We understand (1) as an Itô-stochastic differential equation

$$AX(s)\Big|_{t_0}^{t} + \int_{t_0}^{t} f(s, X(s))ds + \sum_{r=1}^{m}\int_{t_0}^{t} g_r(s, X(s))dW_r(s) = 0\,, \tag{2}$$

where the second integrals are Itô-integrals, and W denotes an m-dimensional Wiener process (or Brownian motion) given on the probability space $(\Omega, \mathcal{F}, P)$ with a filtration $(\mathcal{F}_t)_{t\geq t_0}$. The solution X is a stochastic process depending on the time t and on the random sample $\omega \in \Omega$. Typical paths are nowhere differentiable. In the literature on numerical methods for SDEs (see, e.g., [8, 9, 10, 12]) mainly two concepts of convergence are discussed, weak and strong convergence. Weak convergence relates to Monte-Carlo methods and is mostly concerned with statistical properties of the solutions of SDEs. The term *strong* convergence is often used synonymously for the expression *mean-square* convergence, i.e., convergence in the norm $\|\cdot\|_{L_2}$. Here solution paths for given paths of the driving Wiener process have to be approximated. We denote by $|\cdot|$ the Euclidian norm in $\mathbb{R}^n$, by $\|\cdot\|$ the corresponding induced matrix norm and by $\|Z\|_{L_2} := (I\!E|Z|^2)^{1/2}$ the norm of a vector-valued square-integrable random variable $Z \in L_2(\Omega, \mathbb{R}^n)$.

Only solution paths reveal the phase noise, which is a very important issue in circuit simulation. Subsequently we discuss *mean-square convergence* of numerical schemes for SDAEs of the form (2), where the deterministic part has globally DAE-index 1. A crucial point in designing schemes for SDAEs is to force the iterates to fulfill the constraints at the current time point. That way it is possible to derive schemes for SDAEs from schemes for SDEs ($A = I$) and carry over the

convergence results that are known there. However, terms that include derivatives of the drift and diffusion coefficients f, g_r would lead to terms that also involve the derivative of the solution with respect to the inherent dynamical components and should be avoided.

In the following we discuss schemes that are specially suited for SDEs and SDAEs with small noise. We start with the drift-implicit Euler-scheme

$$A\frac{X_\ell - X_{\ell-1}}{h_\ell} + f(t_\ell, X_\ell) + \sum_{j=1}^{m} g^j(t_{\ell-1}, X_{\ell-1})\frac{\Delta W_\ell^j}{h_\ell} = 0\ , \quad \ell = 1, \ldots, N\ , \tag{3}$$

with a given consistent initial value X_0 , where X_ℓ denotes the approximation to $X(t_\ell)$, $h_\ell = t_\ell - t_{\ell-1}$, $\Delta W_\ell^j = W^j(t_\ell) - W^j(t_{\ell-1}) \sim N(0,h)$ on the deterministic grid $0 = t_0 < t_1 < \ldots < t_N = T$.

We aim at estimates of the mean-square global error $\max_{\ell=0,\ldots,N} \|X(t_\ell) - X_\ell\|_{L_2}$. Also in the stochastic setting there exists a stability inequality by means of which one can estimate the global errors by local ones. If the local error L_ℓ at time-point t_ℓ is defined as the defect that is obtained when the exact solution values are inserted into the numerical scheme, we have, see [1, 16] and, for a related concept of local errors, [10, 12],

$$\max_{\ell=1,\ldots,N} \|X(t_\ell) - X_\ell\|_{L_2} \le S \cdot \max\big(\frac{\|L_\ell\|_{L_2}}{h_\ell^{1/2}}\ ,\ \frac{\|E(L_\ell|\mathcal{F}_{t_{\ell-1}})\|_{L_2}}{h_\ell}\big), \tag{4}$$

with a grid-independent constant S. In general, the order of mean-square convergence of schemes that involve only increments of the driving Wiener process is only 1/2. For additive noise the order of strong convergence becomes 1. However, when the noise is small and the step-sizes are not asymptotically small the error behaviour is still dominated by the deterministic terms. The theoretical order 1/2 of the schemes would be observed only for much smaller stepsizes. Let us express the smallness of the noise by means of a small parameter ϵ in the diffusion coefficient $(g_r(t,x) = \epsilon \cdot \bar{g}_r(t,x)\ r = 1, \ldots, m,\ \epsilon << 1)$ [11]. Then the error of the Euler scheme is bounded by $\mathcal{O}(h + \epsilon^2 h^{1/2})$. It is even worth to use schemes with the deterministic order 2 like the stochastic two-step BDF scheme [1, 2, 14] which show global error $\mathcal{O}(h^2 + \epsilon h + \epsilon^2 h^{1/2})$, or, especially for additive noise, to include mixed classical stochastic integrals of the driving Wiener process into the scheme to get rid of the error terms of order $\mathcal{O}(\epsilon h)$ [3].

Based on the knowledge of mean-square local and global errors we further present an error estimate and, based on this, a stepsize control for the drift-implicit Euler scheme for problems with small noise [13]. This stepsize control leads to adaptive stepsize sequences that are uniform for all paths. Using the local information from a number of simultaneously computed paths, it smoothes the stepsize-sequence and reduces the number of rejected steps.

Using the techniques developed in [15] we aim at an estimate of the local error, and based on this a stepsize control, for schemes with deterministic order 2 in the case of small noise.

References

[1] E. Buckwar and R. Winkler, *Multi-step methods for SDEs and their application to problems with small noise,* SIAM J. Num. Anal., **44:2** (2006), 779–803.

[2] E. Buckwar and R. Winkler, *On two-step schemes for SDEs with small noise,* PAMM **4** (2004), 15–18.

[3] E. Buckwar and R. Winkler, *Improved linear multi-step methods for stochastic ordinary differential equations,* to appear in J. Comput. Appl. Math..

[4] A. Demir, A. Sangiovanni-Vincentelli, *Analysis and simulation of noise in nonlinear electronic circuits and systems,* Kluwer Academic Publishers, 1998.

[5] G. Denk, R. Winkler, *Modeling and simulation of transient noise in circuit simulation,* to appear in: Mathematical and Computer Modelling of Dynamical Systems (MCMDS).

[6] D. Estevez Schwarz, C. Tischendorf, *Structural analysis for electronic circuits and consequences for MNA,* Int. J. Circ. Theor. Appl., **28** (2000), 131–162.

[7] F. Günther, U. Feldmann, *CAD-based electric-circuit modeling in industry I, mathematical structure and index of network equations,* Surv. Math. Ind.,**8** (1999), 97–129.

[8] D. J. Higham, *An algorithmic introduction to numerical simulation of stochastic differential equations,* SIAM Review, **43** (2001), 525–546.

[9] P. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equations,* Springer, Berlin, 1992.

[10] G. Milstein, *Numerical Integration of Stochastic Differential Equations,* Kluwer, 1995, translation from the Russion original of 1988.

[11] G. Milstein, M. Tretyakov, *Mean-square numerical methods for stochastic differential equations with small noise,* SIAM J. Sci. Comput. **18** (1997) 1067–1087.

[12] G. Milstein, M. Tretyakov, *Stochastic Numerics for Mathematical Physics,* Springer Verlag, Berlin, 2004.

[13] W. Römisch and R. Winkler, *Stepsize control for mean-square numerical methods for SDEs with small noise,* SIAM J. Sci. Comput., to appear.

[14] T. Sickenberger, *Mean-square convergence of stochastic multi-step methods with variable step-size,* Preprint 2005-20, Institut für Mathematik, Humboldt-Universität Berlin, 2005.

[15] T. Sickenberger, E. Weinmüller, and R. Winkler, *Local error estimates for moderately smooth ODEs and DAEs,* Preprint 2006-1, Institut für Mathematik, Humboldt-Universität Berlin, 2006. submitted.

[16] R. Winkler, *Stochastic differential algebraic equations of index 1 and applications in circuit simulation,* J. Comput. Appl. Math., **157:2** (2003), 477–505.

Reporter: René Lamour

Participants

Dr. Carmen Arevalo
Dept. of Numerical Analysis
Lund University
Box 118
S-221 00 Lund

Prof. Dr. Martin Arnold
Fachbereich Mathematik u.Informatik
Martin-Luther-Universität
Halle-Wittenberg
06099 Halle

Dr. Andreas Bartel
FB C: Mathematik u. Naturwissensch.
Bergische Universität Wuppertal
42097 Wuppertal

Prof. Dr. Paul Barton
Department of Chemical Engineering
Massachusetts Institute of
Technology
Cambridge, MA 02139
USA

Prof. Dr.Dr.h.c. Hans Georg Bock
Interdisziplinäres Zentrum
für Wissenschaftliches Rechnen
Universität Heidelberg
Im Neuenheimer Feld 368
69120 Heidelberg

Martin Bodestedt
Institut für Mathematik
Technische Universität Berlin
Sekr. MA 4-5
Strasse des 17. Juni 136
10623 Berlin

Dr. Rainer Callies
Zentrum Mathematik
TU München
Boltzmannstr. 3
85748 Garching bei München

Prof. Dr. Stephen L. Campbell
Department of Mathematics
North Carolina State University
Campus Box 8205
Raleigh, NC 27695-8205
USA

Prof. Dr. Elena Celledoni
Dept. of Mathematical Sciences
Norwegian University of Science
and Technology
A. Getz vei 1
N-7491 Trondheim

Dr. Diana Estévez Schwarz
Infineon Technologies
MP TI CS ATS
Balanstr. 73
81541 München

Prof. Dr. Angelo Favini
Dipartimento di Matematica
Universita degli Studi di Bologna
Piazza di Porta S. Donato, 5
I-40126 Bologna

Prof. Dr. Claus Führer
Centre of Mathematical Sciences
Lund University
P.O. Box 118
S-22100 Lund

Prof. Dr. C. William Gear
Princeton University
17, Honey Brook Drive
Princeton, NJ 08540
USA

Prof. Dr. Michael Günther
FB C: Mathematik u. Naturwissensch.
Bergische Universität Wuppertal
42097 Wuppertal

Prof. Dr. Ernst Hairer
Departement de Mathematiques
Universite de Geneve
Case Postale 64
2-4 rue du Lievre
CH-1211 Geneve 4

Dr. Michael Hanke
School of Computer Science and
Communication
Royal Institute of Technology
S-10044 Stockholm

Prof. Dr. Marlis Hochbruck
Mathematisches Institut
Universität Düsseldorf
Universitätsstr. 1
40225 Düsseldorf

Prof. Dr. Achim Ilchmann
Institut f. Mathematik
Technische Universität Ilmenau
Weimarer Str. 25
98693 Ilmenau

Shivakumar Kameswaran
Chemical Engineering Department
Carnegie Mellon University
5000, Forbes Avenue
Pittsburgh PA 15213
USA

Dr. Ekaterina A. Kostina
Interdisziplinäres Zentrum
für Wissenschaftliches Rechnen
Universität Heidelberg
Im Neuenheimer Feld 368
69120 Heidelberg

Prof. Dr. Peter Kunkel
Fakultät für Mathematik/Informatik
Universität Leipzig
Augustusplatz 10/11
04109 Leipzig

Prof. Dr. Galina A. Kurina
Voronezh State Forestry Academy
Ul.Timirjazeva 8
Voronezh, 394613
RUSSIA

Prof. Dr. Anne Kvaerno
Dept. of Mathematical Sciences
Norwegian University of Science
and Technology
A. Getz vei 1
N-7491 Trondheim

Dr. René Lamour
Institut für Mathematik
Math. Naturw. Fakultät II
Humboldt-Universität
10099 Berlin

Dr. Vu Hoang Linh
Faculty of Mathematics, Mechanics
and Informatics
Vietnam National University
334 Nguyen Trai St.
Hanoi
Vietnam

Prof. Dr. Christian Lubich
Mathematisches Institut
Universität Tübingen
Auf der Morgenstelle 10
72076 Tübingen

Dipl.Math. Christoph Lunk
Zentrum Mathematik
TU München
Boltzmannstr. 3
85748 Garching bei München

Prof. Dr. Roswitha März
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin

Prof. Dr. Volker Mehrmann
Institut für Mathematik
Technische Universität Berlin
Sekr. MA 4-5
Strasse des 17. Juni 136
10623 Berlin

Prof. Dr. Alexander Ostermann
Institut für Mathematik
Universität Innsbruck
Technikerstr. 25
A-6020 Innsbruck

Prof. Dr. Linda R. Petzold
Department of Computer Science and
Department of Mechanical and
Environmental Engineering
University of California
Santa Barbara CA 93106-5070
USA

Dr. Roland Pulch
FB C: Mathematik u. Naturwissensch.
Bergische Universität Wuppertal
42097 Wuppertal

Dr. Timo Reis
Institut für Mathematik
Technische Universität Berlin
Sekr. MA 4-5
Strasse des 17. Juni 136
10623 Berlin

Prof. Dr. Peter Rentrop
Zentrum Mathematik
TU München
Boltzmannstr. 3
85748 Garching bei München

Prof. Dr. Werner C. Rheinboldt
Zentrum Mathematik
TU München
Boltzmannstr. 3
85748 Garching bei München

Prof. Dr. Ricardo Riaza
Depto. de Matematica Aplicada a las
Tecnologias de la Informacion
E.T.S.Ingenieros de Telecommunic.
Universidad Politecnica Madrid
E-28040 Madrid

Dr. Monica Selva Soto
Mathematisches Institut
Universität zu Köln
Weyertal 86 - 90
50931 Köln

Prof. Dr. Alla A. Shcheglova
Institute Dynamics of Systems
and Control Theory
P.O. Box 1233
ul. Lermontova 134
664033 Irkutsk
RUSSIA

Prof. Dr. Bernd Simeon
Zentrum Mathematik
TU München
Boltzmannstr. 3
85748 Garching bei München

Prof. Dr. Gustaf Söderlind
Dept. of Numerical Analysis
Lund University
Box 118
S-221 00 Lund

Dr. Tatjana Stykel
Institut für Mathematik
Technische Universität Berlin
Sekr. MA 3-3
Straße des 17. Juni 136
10623 Berlin

Dipl.Math.tech. Gunnar Teichelmann
Zentrum Mathematik
TU München
Boltzmannstr. 3
85748 Garching bei München

Prof. Dr. Caren Tischendorf
Mathematisches Institut
Universität zu Köln
Weyertal 86 - 90
50931 Köln

Prof. Dr. Jan G. Verwer
Centrum voor Wiskunde en
Informatica
Kruislaan 413
NL-1098 SJ Amsterdam

Steffen Voigtmann
Infineon Technologies
MP PD CS TTN
Balanstr.73
81541 München

Prof. Dr. Ewa B. Weinmüller
Institut für Analysis und
Scientific Computing
Technische Universität Wien
Wiedner Hauptstr. 8 - 10
A-1040 Wien

Dr. Renate Winkler
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin

Prof. Dr. Eva Zerz
Lehrstuhl D für Mathematik
RWTH Aachen
Templergraben 64
52062 Aachen