Numerical solution of optimal control problems for descriptor systems

Volker Mehrmann

TU Berlin
Institut für Mathematik

DFG Research Center
MATHEON

Harrachov 23.08.07

joint with Peter Kunkel, U. Leipzig, thanks to DFG and RIP/Oberwolfach

Optimal control of descriptor systems

Optimal control problem:

$$\mathcal{J}(x,u) = \mathcal{M}(x(\bar{t})) + \int_{\underline{t}}^{\bar{t}} \mathcal{K}(t,x,u)\, dt = \min!$$

subject to a descriptor system (differential-algebraic, DAE) constraint

$$F(t,x,u,\dot{x}) = 0, \;\; x(\underline{t}) = \underline{x}.$$

x–state, u–input.

Linear quadratic optimal control

Cost functional:

$$\mathcal{J}(x,u) = \frac{1}{2}x(\overline{t})^T M x(\overline{t}) + \frac{1}{2}\int_{\underline{t}}^{\overline{t}} (x^T W x + 2x^T S u + u^T R u)\, dt,$$

$W = W^T \in C^0(\mathbb{I}, \mathbb{R}^{n,n})$, $S \in C^0(\mathbb{I}, \mathbb{R}^{n,l})$, $R = R^T \in C^0(\mathbb{I}, \mathbb{R}^{l,l})$, $M = M^T \in \mathbb{R}^{n,n}$.

Constraint:

$$E(t)\dot{x} = A(t)x + B(t)u + f, \quad x(\underline{t}) = \underline{x},$$

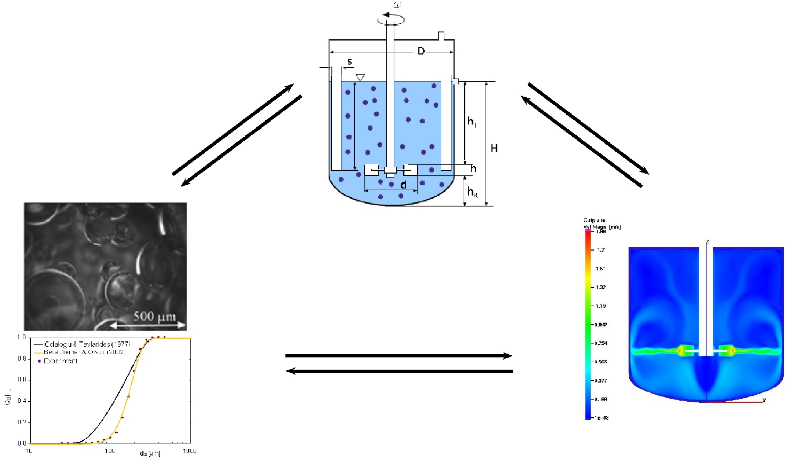
$E \in C^1(\mathbb{I}, \mathbb{R}^{n,n})$, $A \in C^0(\mathbb{I}, \mathbb{R}^{n,n})$, $B \in C^0(\mathbb{I}, \mathbb{R}^{n,l})$, $f \in C^0(\mathbb{I}, \mathbb{R}^{n})$, $\underline{x} \in \mathbb{R}^n$.

Here: Determine optimal controls $u \in \mathbb{U} = C^0(\mathbb{I}, \mathbb{R}^l)$.,

More general spaces, nonsquare and inf. dim. E, A possible.

Drop size distributions in stirred liquid/liquid systems

with M. Kraume from Chemical Engineering (S. Schlauch)



Technological Application, Tasks

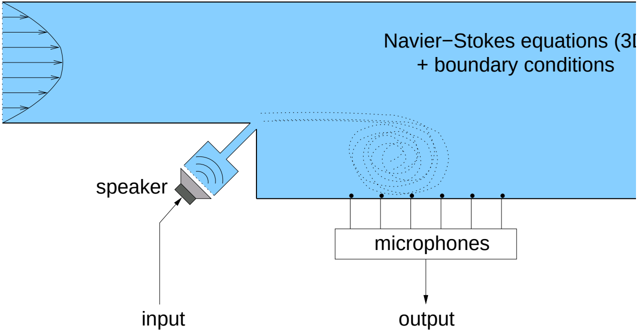
Chemical industry: pearl polymerization and extraction processes

- Modelling of coalescence and breakage in turbulent flow. Navier Stokes equation (flow field), population balance equation (drop size distribution).
- Numerical methods for simulation of coupled system.
- Development of optimal control methods for coupled system.
- Model reduction and observer design.
- Feedback control of real configurations via stirrer speed.

Ultimate goal: Achieve specified average drop diameter and small standard deviation for distribution by real time-control of stirrer-speed.
Space discretization leads to large control system of nonlinear DAEs.

Active flow control, SFB 557

with F. Tröltzsch (M. Schmidt)
Test case (backward step to compare experiment/numerics.)



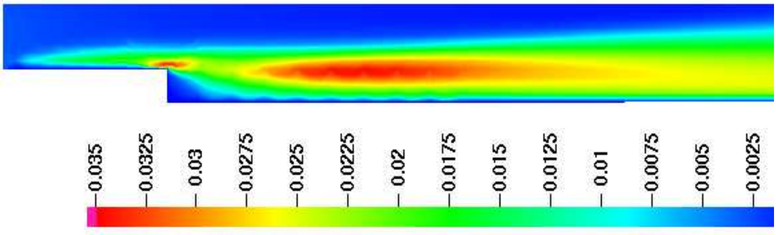
Technological Application, Tasks

Control of detached turbulent flow on airline wing

- Modelling of turbulent flow.
- Development of control methods for large scale systems.
- Model reduction and observer design.
- Optimal feedback control of real configurations via blowing and sucking of air in wing.

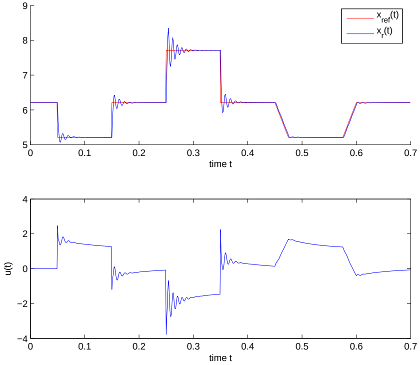
Ultimate goal: Force detached flow back to wing.

Simulated flow



Controlled flow

Movement of recirculation bubble following reference curve.



DAE control systems

After space discretization these problems are DAE control systems

$$F(t, x, \dot{x}, u) = 0,$$

or in the linear case (linearization along solutions)

$$E(t)\dot{x}(t) = A(t)x(t) + B(t)u(t) + f(t),$$

Using a behavior approach, i.e., forming $z(t) = (x, u)$ we obtain general non-square DAEs

$$\mathcal{F}(t, z, \dot{z}) = 0, \quad \mathcal{E}(t)\dot{z} = \mathcal{A}(t)z.$$

The behavior approach allows a uniform mathematical treatment of simulation and control problems!

Why DAEs and not ODEs?

DAEs provide a unified framework for the analysis, simulation and control of coupled dynamical systems (continuous and discrete time).

- Automatic modelling leads to DAEs. (Constraints at interfaces).
- Conservation laws lead to DAEs. (Conservation of mass, energy, volume, momentum).
- Coupling of solvers leads to DAEs (discrete time).
- Control problems are DAEs (behavior).

How does one solve such problems today?

- Simplified models.
- Space discretization with very coarse meshes.
- Identification and realization of black box models.
- Model reduction (mostly based on heuristic methods).
- Coupling of simulation packages.
- Use of standard optimal control techniques for simplified mathematical model.

Future: Solve optimality system for original model

Is there anything to do ?

Why not just apply the Pontryagin maximum principle?

- DAEs (of high index) are often difficult numerically and analytically.
- The (differentiation) index describes the number of differentiations that are needed to turn the problem into an (implicit) ODE (regularity measure).
- For linear ODEs the initial value problem has a unique solution $x \in C^1(\mathbb{I}, \mathbb{R}^n)$ for every $u \in \mathbb{U}$, every $f \in C^0(\mathbb{I}, \mathbb{R}^n)$, and every initial value $\underline{x} \in \mathbb{R}^n$.
- DAEs, where $E(t)$ is singular, may not be (uniquely) solvable for all $u \in \mathbb{U}$ and the initial conditons are restricted.

Previous work: (general case open, since the 70ies)

- Linear constant coefficient index 1 case, Bender/Laub 87, Campbell 87 M. 91, Geerts 93.
- Regularization to index 1, Bunse-Gerstner/M./Nichols 92, 94, Byers/Geerts/M. 97, Byers/Kunkel/M. 97.
- Linear variable coefficients index 1 case, Kunkel./M. 97.
- Semi-explicit nonlinear index 1 case, maximum principle, De Pinho/Vinter 97, Devdariani/Ledyaev 99.
- Semi-explicit index $2, 3$ case Roubicek/Valasek 02.
- Linear index $1, 2$ case with properly stated leading term, Balla/März, 02,04, Balla/Linh 05, Kurina/März 04, Backes 06.
- Multibody systems (structured and of index 3), Büskens/Gerdts 00, Gerdts 03,06.

A crash course in DAE Theory

For the numerical solution of general DAEs and for the design of controllers, we use derivative arrays (Campbell 1989).

We assume that derivatives of original functions are available or can be obtained via computer algebra or automatic differentiation.

Linear case: We put $E(t)\dot{x} = A(t)x + f(t)$ and its derivatives up to order μ into a large DAE

$$M_k(t)\dot{z}_k = N_k(t)z_k + g_k(t), \quad k \in \mathbb{N}_0$$

for $z_k = (x, \dot{x}, \ldots, x^{(k)})$.

$$M_2 = \begin{bmatrix} E & 0 & 0 \\ A - \dot{E} & E & 0 \\ \dot{A} - 2\ddot{E} & A - \dot{E} & E \end{bmatrix}, \; N_2 = \begin{bmatrix} A & 0 & 0 \\ \dot{A} & 0 & 0 \\ \ddot{A} & 0 & 0 \end{bmatrix}, \; z_2 = \begin{bmatrix} x \\ \dot{x} \\ \ddot{x} \end{bmatrix}.$$

Theorem, Kunkel/M. 1996 Under some constant rank assumptions, for a linear DAE there exist integers μ, a, d and v such that:

1. $\operatorname{corank} M_{\mu+1}(t) - \operatorname{corank} M_{\mu}(t) = v$.
2. $\operatorname{rank} M_{\mu}(t) = (\mu+1)m - a - v$ on $\mathbb{I}$, and there exists a smooth matrix function $Z_{2,3}$ (left nullspace of M_μ) with $Z_{2,3}^T M_\mu(t) = 0$.
3. The projection $Z_{2,3}$ can be partitioned into two parts: Z_2 (left nullspace of $[M_\mu, N_\mu]$) so that the first block column $\hat{A}_2$ of $Z_2^* N_\mu(t)$ has full rank a and $Z_3^* N_\mu(t) = 0$. Let T_2 be a smooth matrix function such that $\hat{A}_2 T_2 = 0$, (right nullspace of $\hat{A}_2$).
4. $\operatorname{rank} E(t) T_2 = d = l - a - v$ and there exists a smooth matrix function Z_1 of size (n, d) with $\operatorname{rank} \hat{E}_1 = d$, where $\hat{E}_1 = Z_1^T E$.

Reduced problem

- The quantity μ is called the strangeness-index. It describes the smoothness requirements for forcing or input functions.
- It generalizes the d-index to over- and underdetermined DAEs (and counts differently).
- We obtain a **numerically computable** reduced system:

$$\begin{array}{rcll} \hat{E}_1(t)\dot{x} & = & \hat{A}_1(t)x + \hat{f}_1(t), & d \text{ differential equations} \\ 0 & = & \hat{A}_2(t)x + \hat{f}_2(t), & a \text{ algebraic equations} \\ 0 & = & \hat{f}_3(t), & v \text{ consistency equations} \end{array}$$

where $\hat{A}_1 = Z_1^T A$, $\hat{f}_1 = Z_1^T f$, and $\hat{f}_2 = Z_2^T g_\mu$, $\hat{f}_3 = Z_3^T g_\mu$.

- The reduced system has the same solution set as the orignal problem but now it has strangeness-index 0. Remodeling!
- **We assume from now on that we have the reduced system.**

Calculus of variations for linear ODEs (E=I)

Introduce Lagrange multiplier function $\lambda(t)$ and couple constraint into cost function, i.e. minimize

$$\begin{aligned}\tilde{\mathcal{J}}(x,u,\lambda) &= \frac{1}{2}x(\bar{t})^T M x(\bar{t}) + \frac{1}{2}\int_{\underline{t}}^{\bar{t}} (x^T W x + 2x^T S u + u^T R u) \\ &+ \lambda^T(\dot{x} - Ax + Bu + f)\, dt.\end{aligned}$$

Consider $x + \delta x$, $u + \delta u$ and $\lambda + \delta\lambda$.
For a minimum the cost function has to go up in the neighborhood, so we get optimality conditions (Euler-Lagrange equations):

Optimality system

Theorem If (x, u) is a solution to the optimal control problem, then there exists a Lagrange multiplier function $\lambda \in C^1(\mathbb{I}, \mathbb{R}^n)$, such that (x, λ, u) satisfy the *optimality boundary value problem*

$$
\begin{aligned}
&\text{(a)} \quad \dot{x} = Ax + Bu + f, \; x(\underline{t}) = \underline{x},\\
&\text{(b)} \quad \dot{\lambda} = Wx + Su - A^T\lambda, \; \lambda(\bar{t}) = -Mx(\bar{t}),\\
&\text{(c)} \quad 0 = S^T x + Ru - B^T\lambda.
\end{aligned}
$$

The adjoint equation always has a unique solution.

Naive Idea for DAEs

Replace the identity in front of $\dot{x}$ by $E(t)$ and then do the analysis in the same way.
For DAEs the formal optimality system then could be

$$
\begin{array}{lrl}
\text{(a)} & E\dot{x} & = Ax + Bu + f,\ x(\underline{t}) = \underline{x} \\
\text{(b)} & \frac{d}{dt}(E^T\lambda) & = Wx + Su - A^T\lambda,\ (E^T\lambda)(\bar{t}) = -Mx(\bar{t}), \\
\text{(b)} & 0 & = S^T x + Ru - B^T\lambda.
\end{array}
$$

This works if the system has strangeness-index $\mu = 0$ as a free system with $u = 0$ but not in general.

What are the difficulties ?

- In the proof one has to guarantee that the resulting adjoint equation for λ has a unique solution.
- But in the DAE case the formal adjoint equation may not have a (unique) solution.
- The formal boundary conditions may not be consistent.
- The solution of the optimality system may not exist or may not be unique.

Example

Consider

$$\mathcal{J}(x,u) = \frac{1}{2}\int_0^1 (x_1^2 + u^2)dt = \min!$$

subject to the differential-algebraic system

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u + \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}.$$

A simple calculation yields the optimal solution

$$x_1 = u = \lambda_1 = -\tfrac{1}{2}(f_1 + \dot{f}_2), \quad x_2 = -f_2, \quad \lambda_2 = 0.$$

In the formal optimality system we get

$$x_1 = u = \lambda_1 = -\frac{1}{2}(f_1 + \dot{f}_2), \quad x_2 = -f_2, \quad \lambda_2 = -\frac{1}{2}(\dot{f}_1 + \ddot{f}_2)$$

without using the initial condition $\lambda_1(1) = 0$.

- The formal initial condition may be consistent or not. This initial condition should not be present.
- Moreover, λ_2 requires more smoothness of the inhomogeneity than in the optimal solution.

Further examples, see Dissertation Backes 06.

Solution space

To derive optimality conditions for DAEs, we need the right solution space for x.

$$\mathbb{X} = C^1_{E^+E}(\mathbb{I}, \mathbb{R}^n) = \left\{ x \in C^0(\mathbb{I}, \mathbb{R}^n) \mid E^+Ex \in C^1(\mathbb{I}, \mathbb{R}^n) \right\},$$

where E^+ denotes the Moore-Penrose inverse of the matrix valued function $E(t)$, i.e. the unique matrix function that satisfies the Penrose axioms.

$$EE^+E = E, \quad E^+EE^+ = E^+, \quad (EE^+)^T = EE^+, \quad (E^+E)^T = E^+E$$

The input space $\mathbb{U}$ is usually a set of piecewise continuous functions or a space of distributions.

Necessary optimality condition

Theorem Consider the linear quadratic DAE optimal control problem with a consistent initial condition. Suppose that the system has $\mu = 0$ as a behavior system and that $Mx(\bar{t}) \in \operatorname{cokernel} E(\bar{t})$.
If $(x, u) \in \mathbb{X} \times \mathbb{U}$ is a solution to this optimal control problem, then there exists a Lagrange multiplier function $\lambda \in C^1_{E^+E}(\mathbb{I}, \mathbb{R}^n)$, such that (x, λ, u) satisfy the optimality boundary value problem

$$E\frac{d}{dt}(E^+Ex) = (A + E\frac{d}{dt}(E^+E))x + Bu + f,\ (E^+Ex)(\underline{t}) = \underline{x},$$
$$E^T\frac{d}{dt}(EE^+\lambda) = Wx + Su - (A + EE^+\dot{E})^T\lambda,\ (EE^+\lambda)(\bar{t}) = -E^+(\bar{t})^T Mx(\bar{t}),$$
$$0 = S^Tx + Ru - B^T\lambda.$$

Sufficient condition

Theorem Consider the optimal control problem with a consistent initial condition and suppose that in the cost functional we have that

$$\begin{bmatrix} W & S \\ S^T & R \end{bmatrix}, \; M$$

are (pointwise) positive semidefinite. If (x^*, u^*, λ) satisfies the (formal) optimality system then for any (x, u) satisfying the constraint we have

$$\mathcal{J}(x, u) \geq \mathcal{J}(x^*, u^*).$$

Remarks

- If a minimum exists, then it satisfies the optimality system.
- If a unique solution to the formal optimality system exists, then x, u are the same as from the optimality system, λ may be different.
- The optimality DAE may have $\mu > 0$. Then it is numerically difficult to solve and further consistency conditions or smoothness requirements arise.
- The condition that the original system has $\mu = 0$ as a behavior system is not necessary if the cost function is chosen appropriately, so that the resulting optimality system has $\mu = 0$.

Differential-algebraic Riccati equations

If R in the cost functional is invertible, and if the system has $\mu = 0$ as a free system with $u = 0$, then one can (at least in theory) apply the usual Riccati approach to

$$
\begin{aligned}
&E\tfrac{d}{dt}(E^+Ex) = (A + E\tfrac{d}{dt}(E^+E))x + Bu + f,\ (E^+Ex)(\underline{t}) = \underline{x},\\
&E^T\tfrac{d}{dt}(EE^+\lambda) = Wx + Su - (A + EE^+\dot{E})^T\lambda,\ (EE^+\lambda)(\bar{t}) = -E^+(\bar{t})^T Mx(\bar{t}),\\
&0 = S^Tx + Ru - B^T\lambda.
\end{aligned}
$$

If $\mu > 0$ or R is singular, then the Riccati approach may not work, even if the boundary value problem has a unique solution.

Nonlinear problems

For nonlinear systems $F(t, x, \dot{x}) = 0$ one considers nonlinear derivative arrays:

$$0 = F_k(t, x, \dot{x}, \ldots, x^{(k+1)}) = \begin{bmatrix} F(t, x, \dot{x}) \\ \frac{d}{dt} F(t, x, \dot{x}) \\ \ldots \\ \frac{d^k}{dt^k} F(t, x, \dot{x}) \end{bmatrix}.$$

We set

$$\begin{aligned} M_k(t, x, \dot{x}, \ldots, x^{(k+1)}) &= F_{k;\dot{x},\ldots,x^{(k+1)}}(t, x, \dot{x}, \ldots, x^{(k+1)}), \\ N_k(t, x, \dot{x}, \ldots, x^{(k+1)}) &= -(F_{k;x}(t, x, \dot{x}, \ldots, x^{(k+1)}), 0, \ldots, 0), \\ z_k &= (t, x, \dot{x}, \ldots, x^{(k+1)}). \end{aligned}$$

Hypothesis: There exist integers μ, r, a, d, and v such that $\mathbf{L} = F_\mu^{-1}(\{0\}) \neq \emptyset$.
We have $\operatorname{rank} F_{\mu;t,x,\dot{x},\ldots,x^{(\mu+1)}} = \operatorname{rank} F_{\mu;x,\dot{x},\ldots,x^{(\mu+1)}} = r$, in a neighborhood of $\mathbf{L}$ such that there exists an equivalent system $\tilde{F}(z_\mu) = 0$ with a Jacobian of full row rank r. On $\mathbf{L}$ we have

1. $\operatorname{corank} F_{\mu;x,\dot{x},\ldots,x^{(\mu+1)}} - \operatorname{corank} F_{\mu-1;x,\dot{x},\ldots,x^{(\mu+1)}} = v$.
2. $\operatorname{corank} \tilde{F}_{x,\dot{x},\ldots,x^{(\mu+1)}} = a$ and there exist smooth matrix functions Z_2 (left nullspace of M_μ) and T_2 (right nullspace of $\hat{A}_2 = \tilde{F}_x$) with $Z_2^T \tilde{F}_{x,\dot{x},\ldots,x^{(\mu+1)}} = 0$ and $Z_2^T \hat{A}_2 T_2 = 0$.
3. $\operatorname{rank} F_{\dot{x}} T_2 = d, \quad d = m - a - v$, and there exists a smooth matrix function Z_1 with $\operatorname{rank} Z_1^T F_{\dot{x}} = d$.

Theorem Kunkel/M. 2002 The solution set **L** forms a (smooth) manifold of dimension $(\mu+2)n+1-r$.
The DAE can locally be transformed (by application of the implicit function theorem) to a reduced DAE of the form

$$
\begin{aligned}
\dot{x}_1 &= G_1(t, x_1, x_3), && (d \text{ differential equations}),\\
x_2 &= G_2(t, x_1, x_3), && (a \text{ algebraic equations}),\\
0 &= 0 && (v \text{ redundant equations}).
\end{aligned}
$$

The variables x_3 represent undetermined components (controls).

Optimality conditions

Assume that $\mu = 0$ for the system in behavior form with $z = (x, u)$, then in terms of the reduced DAE, the local optimality system is

$$
\begin{array}{ll}
\text{(a)} & \dot{x}_1 = \mathcal{L}(t, x_1, u),\ x_1(\underline{t}) = \underline{x}_1, \\
\text{(b)} & x_2 = \mathcal{R}(t, x_1, u), \\
\text{(c)} & \dot{\lambda}_1 = \mathcal{K}_{x_1}(t, x_1, x_2, u)^T - \mathcal{L}_{x_1}(t, x_1, x_2, u)^T \lambda_1 - \mathcal{R}_{x_1}(t, x_1, u)^T \lambda_1, \\
& \qquad \lambda_1(\bar{t}) = -\mathcal{M}_{x_1}(x_1(\bar{t}), x_2(\bar{t}))^T \\
\text{(d)} & 0 = \mathcal{K}_{x_2}(t, x_1, x_2, u)^T + \lambda_2, \\
\text{(e)} & 0 = \mathcal{K}_u(t, x_1, x_2, u)^T - \mathcal{L}_u(t, x_1, u)^T \lambda_1 - \mathcal{R}_u(t, x_1, u)^T \lambda_2, \\
\text{(f)} & \gamma = \lambda_1(\underline{t})
\end{array}
$$

Here λ_1, λ_2 are Lagrange multipliers associated with x_1, x_2 and γ is associated with the initial value constraint.

Remarks

- These are local results.
- All the results can be generalized to general nonsquare nonlinear systems.
- End point conditions for x can be included.
- Input and state constraints can be included to give a maximum principle.

Numerical Methods

Linear case: Given $E(t), A(t), B(t), f(t)$ in the DAE and $S(t), R(t), W(t), M$ from the cost functional.
The resulting linear optimality system has the form

$$
\begin{aligned}
&\text{(a)} \quad \hat{E}_1 \dot{x} = \hat{A}_1 x + \hat{B}_1 u + \hat{f}_1, \;\; (\hat{E}_1^+ \hat{E}_1 x)(\underline{t}) = \underline{x} \\
&\text{(b)} \quad 0 = \hat{A}_2 x + \hat{B}_2 u + \hat{f}_2, \\
&\text{(c)} \quad \tfrac{d}{dt}(\hat{E}_1^T \lambda_1) = Wx + Su - \hat{A}_1^T \lambda_1 - \hat{A}_2^T \lambda_2, \\
&\qquad\qquad \lambda_1(\bar{t}) = -[\, \hat{E}_1^+(\bar{t})^T \;\; 0 \,] M x(\bar{t}), \\
&\text{(d)} \quad 0 = S^T x + Ru - \hat{B}_1^T \lambda_1 - \hat{B}_2^T \lambda_2.
\end{aligned}
$$

where $\hat{E}_i, \hat{A}_i, \hat{B}_i, \hat{f}_i$ are obtained by projection with smooth orthogonal projections Z_i from the derivative array.
An analogous structure arises locally in the nonlinear case.

Numerical Problems

- In the implementation of our numerical integration codes we use nonsmooth projectors Z_1^T, Z_2^T, since it would be too expensive to carry smooth projectors along.
- For numerical forward (in time) simulation, it is enough that we know the existence of smooth projectors.
- Integration methods like Runge-Kutta or BDF do not see the nonsmooth behavior.
- But the adjoint variables (Lagrange multipliers) depend on these projections and their derivatives.

However, even if Z_1^T, Z_2^T are nonsmooth, $Z_1 Z_1^T$ and $Z_2 Z_2^T$ are smooth.

Smooth optimality system

- Choose

$$\hat{E}_1^T\lambda_1 = E^T Z_1\lambda_1 = E^T Z_1 Z_1^T Z_1\lambda_1 = E^T Z_1 Z_1^T\hat{\lambda}_1.$$

- With $\hat{\lambda}_1 = Z_1\lambda_1$ we obtain smooth coefficients for $\hat{\lambda}_1$.
- However, we have to add the condition that $\hat{\lambda}_1 \in \mathrm{range}\, Z_1$ to the system.
- If Z_i' completes Z_i to a full orthogonal matrix (we compute these anyway when doing a QR or SVD computation) then these conditions can be expressed as

$$Z_i'^T\hat{\lambda}_i = 0,\ i = 1, 2$$

New linear optimality system

For the numerical solution we use the optimality system.

(a) $\hat{E}_1 \dot{x} = \hat{A}_1 x + \hat{B}_1 u + \hat{f}_1, \ (\hat{E}_1^+ \hat{E}_1 x)(\underline{t}) = \underline{x},$

(b) $0 = \hat{A}_2 x + \hat{B}_2 u + \hat{f}_2,$

(c) $\frac{d}{dt}(E^T Z_1 Z_1^T \hat{\lambda}_1) = Wx + Su - A^T \hat{\lambda}_1 - [\, I_n \;\; 0 \,|\, 0 \;\; 0 \,|\, \cdots \,|\, 0 \;\; 0 \,] N_\mu^T \hat{\lambda}_2,$
$(Z_1^T \hat{\lambda}_1)(\bar{t}) = -[\, \hat{E}_1^+(\bar{t})^T \;\; 0 \,] M x(\bar{t}),$

(d) $0 = S^T x + Ru - B^T \hat{\lambda}_1 - [\, 0 \;\; I_l \,|\, 0 \;\; 0 \,|\, \cdots \,|\, 0 \;\; 0 \,] N_\mu^T \hat{\lambda}_2$

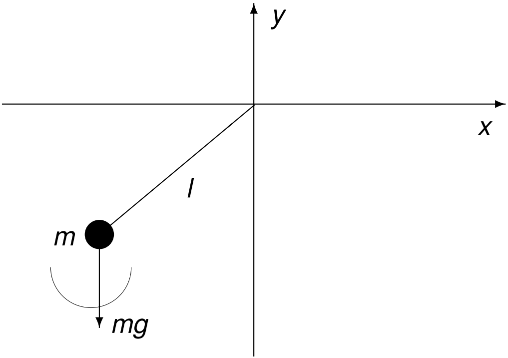
(e) $0 = {Z_1'}^T \hat{\lambda}_1,$

(f) $0 = {Z_2'}^T \hat{\lambda}_2.$

All quantities are available for all time steps.
An analoguous system can be derived for each Gauss-Newton step in the nonlinear case.

Numerical Example

A motor controlled pendulum with a motor in the origin shall be driven into its equilibrium with minimal costs, ex. from Büskens/Gerdts 2002.



Model problem

$$J(x,u) = \int_0^3 u(t)^2 \, dt = \min!$$

s.t.

$$\begin{aligned} \dot{x}_1 &= x_3, & x_1(0) &= \tfrac{1}{2}\sqrt{2}, & g &= 9.81 \\ \dot{x}_2 &= x_4, & x_2(0) &= -\tfrac{1}{2}\sqrt{2}, \\ \dot{x}_3 &= -2x_1x_5 + x_2u, & x_3(0) &= 0, \\ \dot{x}_4 &= -g - 2x_2x_5 - x_1u, & x_4(0) &= 0, \\ 0 &= x_1^2 + x_2^2 - 1, & x_5(0) &= -\tfrac{1}{2}gx_2(0). \end{aligned}$$

- DAE satisfies Hypothesis with $\mu = 2$, $a = 3$, $d = 2$, and $v = 0$.
- Discretization with our DAE/BVP solver (Kunkel/M./Stöver 2004) using midpoint rule for algebraic and trapezoidal rule for differential part, constant stepsize $h = .02$.

Gauss-Newton results

- Tolerance for the Gauß-Newton method was 10^{-7}.
- Let k count the iterations and Δw_k denote the Gauß-Newton correction.

k	$\|\Delta w_k\|_2$
1	0.140D+03
⋮	⋮
17	0.103D+01
18	0.610D-02
19	0.318D-06
20	0.966D-11

- Initial bad convergence is due to a bad initial guess.
- Final value of cost function is $J_{\text{opt}} = 3.82$ which is correct up to discretization and roundoff errors.

Conclusions

- Theoretical analysis (solvability) for general over- and under-determined linear and nonlinear DAEs of arbitrary index.
- Optimality conditions (linear and nonlinear) and maximum principle for general DAEs.
- Model verification, model reduction and removal of redundancies is possible in a numerically stable way.
- Numerical software for linear and nonlinear initial and boundary value problems for DAEs.
- Recent text book. P. Kunkel and V. Mehrmann, Differential algebraic equations. Analysis and numerical solution. European Mathematical Society, Zürich, 2006.

Thank you very much
for your attention.