Quadrature-defect-based a-posteriori error estimates for differential-algebraic equations

W. Auzinger, H. Lehner, E. Weinmüller

Institute for Analysis and Scientific Computing
Vienna University of Technology

Austrian Numerical Analysis Day Linz 2008

DAE with properly stated leading term

Problem class

Consider a DAE with properly stated leading term (according to Balla, März 2002) of tractability index 1

$$\mathbf{A}(t)(\mathbf{D}(t)\mathbf{x}(t))' + \mathbf{B}(t)\mathbf{x}(t) \;=\; \mathbf{g}(t), t \in [0,1],$$

posed as IVP, where

$\mathbf{A}(t) \in \mathbb{R}^{m\times n}$, $\mathbf{D}(t) \in \mathbb{R}^{n\times m}$, $\mathbf{B}(t) \in \mathbb{R}^{m\times m}$, $n \leq m$,
$\mathbf{x}(t)$, $\mathbf{g}(t) \;\in\; \mathbb{R}^m$
$\ker \mathbf{A}(t) = \{0\}$, $\operatorname{im} \mathbf{D}(t) = \mathbb{R}^n$, all functions sufficiently smooth.

DAE with properly stated leading term

Extended system

Assumption $D(t) \equiv \mathrm{const}$ is not restrictive, because

- such a system can be rewritten as

$$\begin{aligned} \mathbf{A}(t)\mathbf{u}'(t) + \mathbf{B}(t)\mathbf{x}(t) &= \mathbf{g}(t) \\ \mathbf{u}(t) - \mathbf{D}(t)\mathbf{x}(t) &= 0 \end{aligned}$$

- or

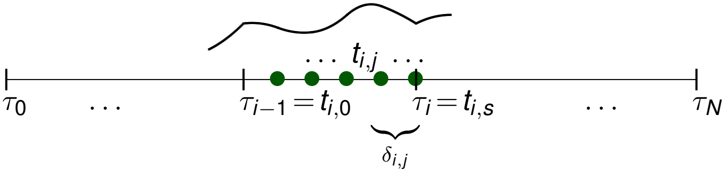
$$\underbrace{\begin{pmatrix} \mathbf{A}(t) \\ 0 \end{pmatrix}}_{A(t)} \left[\underbrace{\begin{pmatrix} 0 & I \end{pmatrix}}_{D} \underbrace{\begin{pmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \end{pmatrix}}_{x(t)} \right]' + \underbrace{\begin{pmatrix} \mathbf{B}(t) & 0 \\ \mathbf{D}(t) & -I \end{pmatrix}}_{B(t)} \underbrace{\begin{pmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \end{pmatrix}}_{x(t)} = \underbrace{\begin{pmatrix} \mathbf{g}(t) \\ 0 \end{pmatrix}}_{g(t)}$$

Method class
Collocation

We consider collocation methods for the numerical solution of such DAEs, i. e.

$$\hat{F}(x) := A(t_{i,j})(Dx(t_{i,j}))' + B(t_{i,j})x(t_{i,j}) - g(t_{i,j}),$$
$$i = 1, \ldots, N, \quad j = 1, \ldots, s$$

for continuous piecewise polynomial functions x of degree $\leq s$, s even.



Abstract setting (nonlinear)

Original problem, working scheme, auxiliary scheme

Consider

- $F^*(x) = 0$ ... original problem, solution x^*
- $\hat{F}(x) = 0$ ... working scheme, solution $\hat{x}$
- $F^{\rm def}(x)$ ... defect-defining operator, typically restriction of F^* to discrete grid; this operator is not inverted numerically
- $\tilde{F}(x) = 0$ ... auxiliary scheme, solution $\tilde{x}$
  $\tilde{F}$ is assumed to be 'cheap to solve', plays auxiliary role in error estimation

  ($F^* \approx F^{\rm def} \approx \hat{F} \approx \tilde{F}$)
- $\hat{x}$ is computed by solving $\hat{F}(x) = 0$

  ... wish to estimate the (global) error $\hat{e} := \hat{x} - x^*$

Defect-based a-posteriori error estimation

DeC approach: Estimate global error using auxiliary scheme

Basic idea due to Zadunaisky 1976, Stetter 1978:

To estimate $\hat{e} = \hat{x} - x^*$, proceed as follows

- Compute defect (residual) $\hat{d} := F^{\mathrm{def}}(\hat{x})$
- Solve $\tilde{F}(x) = 0 \longrightarrow \tilde{x}$
- Solve $\tilde{F}(x) = \hat{d} \longrightarrow \tilde{x}^{\mathrm{def}}$
- Estimate $\hat{e}$:

$$\hat{e} = \hat{x} - x^* \approx F^{*\,-1}\underbrace{F^{\mathrm{def}}(\hat{x})}_{=\,\hat{d}} - F^{*\,-1}\underbrace{F^{\mathrm{def}}(x^*)}_{\approx\,0}$$

$$\approx \tilde{F}^{-1}(\hat{d}) - \tilde{F}^{-1}(0) = \tilde{x}^{\mathrm{def}} - \tilde{x}$$

- I.e.: error estimate $\hat{\epsilon} := \tilde{x}^{\mathrm{def}} - \tilde{x} \approx \hat{x} - x^* =$ error

Auxiliary scheme and pointwise defect

$\tilde{F}$ = backward Euler-like

$\tilde{F}(x) = 0$ ... low order discretization scheme

In particular: Consider backward Euler-like scheme over collocation nodes,

$$A(t_{i,j})\frac{D\tilde{x}_{i,j} - D\tilde{x}_{i,j-1}}{\delta_{i,j}} + B(t_{i,j})\tilde{x}_{i,j} - g(t_{i,j}) \;=\; 0,$$
$$i = 1,\dots,N,\; j = 1,\dots,s$$

$$A(t_{i,j})\frac{D\tilde{x}^{\mathrm{def}}{}_{i,j} - D\tilde{x}^{\mathrm{def}}{}_{i,j-1}}{\delta_{i,j}} + B(t_{i,j})\tilde{x}^{\mathrm{def}}_{i,j} - g(t_{i,j}) \;=\; \hat{d}_{i,j}$$
$$i = 1,\dots,N,\; j = 1,\dots,s$$

Defect-based a-posteriori error estimation

Choice of F^{def}

Choice of F^{def}

- Define point-wise defect

$$\hat{d}^{\circ}(t) := A(t)D\hat{x}'(t) + B(t)\hat{x}(t) - g(t)$$

$\hat{d}^{\circ}_{i,j} = 0$ at all points evaluated in $\tilde{F}(\hat{d}^{\circ}) \Rightarrow$ estimate would always be zero.

- Use modified defect or quadrature defect, therefore QDeC (Auzinger, Koch, Weinmüller 2002 in ODE context)

$$F^{\mathrm{def}}(\hat{x})(t_{i,j}) := \hat{d}_{i,j} := \sum_{k=0}^{s} \alpha_{jk}\hat{d}^{\circ}(t_{i,k}) = \frac{1}{\delta_{i,j}}\int_{t_{i,j-1}}^{t_{i,j}} \hat{d}^{\circ}(t)dt + \mathcal{O}(h^{s+1})$$

In this sum $\alpha_{j,0}\hat{d}^{\circ}(t_{i,0})$ is only non-zero term!

Defect-based a-posteriori error estimation
Main result

Main result

$$\hat{e} = \hat{x} - x^* = \mathcal{O}(h^s)$$

$\hat{\epsilon} = \tilde{x}^{\mathrm{def}} - \tilde{x}$ is an asymptotically correct estimate for the error $\hat{e} = \hat{x} - x^*$, i. e.

$$\hat{\epsilon} - \hat{e} = \mathcal{O}(h^{s+1}).$$

Numerical example
Problem

- Initial value problem

$$\begin{pmatrix} e^t \\ e^t \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} x'(t) + \begin{pmatrix} e^t(1+\cos^2 t) & \cos^2 t \\ e^t(-1+\cos^2 t) & -\cos^2 t) \end{pmatrix} x(t) = \begin{pmatrix} \sin^2 t(1-\cos t) - \sin t \\ \sin^2 t(-1-\cos t) - \sin t \end{pmatrix},$$

  on $[0,1]$ with initial condition $x(0) = (1,-1)^T$.

- We use collocation at equidistant points with $s = 4$ on $N = 2,4,8,16,32$ intervals.
  Asymptotical order $\hat{\epsilon} - \hat{e} = \mathcal{O}(h^{s+1})$ is clearly visible.

Numerical example

Results tabular, component 1

- First component, at $t = 1$:

N	$\hat{e}$	$\mathrm{ord}_{\hat{e}}$	$\hat{\epsilon} - \hat{e}$	$\mathrm{ord}_{\hat{\epsilon}-\hat{e}}$
4	-2.466e−06	3.8	8.513e−08	4.6
8	-1.634e−07	3.9	2.989e−09	4.8
16	-1.051e−08	4.0	9.886e−11	4.9
32	-6.664e−10	4.0	3.180e−12	5.0

- First component, maximum absolute values over all collocation points $\in [0, 1]$:

N	$\hat{e}$	$\mathrm{ord}_{\hat{e}}$	$\hat{\epsilon} - \hat{e}$	$\mathrm{ord}_{\hat{\epsilon}-\hat{e}}$
4	2.732e−06	4.0	1.272e−07	5.3
8	1.711e−07	4.0	3.578e−09	5.2
16	1.074e−08	4.0	1.074e−10	5.1
32	6.734e−10	4.0	3.311e−12	5.0

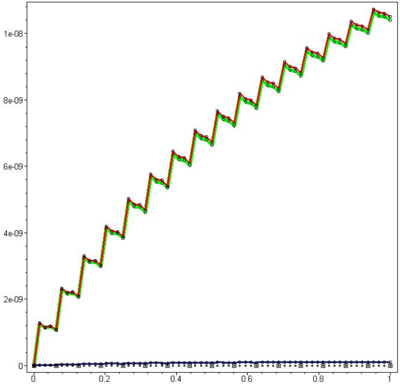
Numerical example

Results tabular, component 2

- Second component, at $t = 1$:

N	$\hat{e}$	$\mathrm{ord}_{\hat{e}}$	$\hat{\epsilon} - \hat{e}$	$\mathrm{ord}_{\hat{\epsilon}-\hat{e}}$
4	2.906e−05	3.8	-7.927e−07	4.6
8	1.522e−06	3.9	-2.783e−08	4.8
16	9.788e−08	4.0	-9.206e−10	4.9
32	6.205e−09	4.0	-2.961e−12	5.0

Numerical example
Results graphical



error $|\hat{e}_1(t)| = |\hat{x}_1(t) - x^*{}_1(t)|$

error estimate $|\hat{\epsilon}_1(t)|$

deviation of error estimate $|\hat{\epsilon}_1(t) - \hat{e}_1(t)|$

Other examples

Other numerical experiments indicate that

- singularities of the first kind do not compromise asymptotical correctness (theoretically established in the ODE case)
- with use of interior collocation points asymptotical correctness is lost, but estimate still gives reasonable approximation of error (preliminary)

Decoupling of Index 1 DAEs
Definition of matrices

Index 1 DAEs with constant $\operatorname{im} D(t)$ can be decoupled into a pure ODE (inherent ODE) and a purely algebraic equation:

- Let Q be a linear projector of $\mathbb{R}^n$ onto $\ker D$,
- D^- a generalized reflexive inverse of D such that

$$D^- D = I - Q, \quad DD^- = I, \quad DD^- D = D, \quad D^- DD^- = D^-$$

- $G(t) = A(t)D + B(t)Q, \quad (\exists G^{-1} \Leftrightarrow \mathit{index} = 1)$
- $N(t) = G(t)^{-1} B(t) D^-$.

Decoupling of Index 1 DAEs
Inherent ODE

- Premultiplying our system with $DG(t)^{-1}$ and $QG(t)^{-1}$, respectively, we obtain

$$\begin{aligned} Dx^{*\prime}(t)) + DN(t)(Dx^*(t)) - DG(t)^{-1}g(t) &= 0, \\ Qx^*(t) + QN(t)(Dx^*(t)) - QG(t)^{-1}g(t) &= 0, \end{aligned}$$

  i. e. the inherent ODE for $Dx(t)$ and an algebraic equation expressing $Qx(t)$ in terms of $Dx(t)$, with

-

$$x^*(t) = (I - Q)x^*(t) + Qx^*(t) = D^- Dx^*(t) + Qx^*(t).$$

Decoupling of Index 1 DAEs

Decoupled collocation and error

- Collocation equations decouple in exactly the same way:

$$\begin{aligned} D\hat{x}'(t_{i,j})) + DN(t_{i,j})\, D\hat{x}(t_{i,j}) - DG(t_{i,j})^{-1} g(t_{i,j}) &= 0, \\ Q\hat{x}(t_{i,j}) + QN(t_{i,j})\, D\hat{x}(t_{i,j}) - QG(t_{i,j})^{-1} g(t_{i,j}) &= 0, \end{aligned}$$

$$i = 1, \ldots, N, \quad j = 1, \ldots, s$$

- Theory of collocation for ODEs and polynomial interpolation argument:
  $D\hat{e}(t),\ D\hat{e}'(t), Q\hat{e}(t), \hat{e}(t) = \mathcal{O}(h^s)$,
- and $\hat{d}^\circ(t_{i,j}) = \mathcal{O}(h^s) \;\Rightarrow\; \hat{d}(t_{i,j}) = \mathcal{O}(h^s)$.

Analysis of deviation

Decoupling of defect definition

- Pointwise defect

$$DG(t)^{-1}\hat{d}^{\circ}(t) = D\hat{x}'(t) + DN(t)D\hat{x}(t) - DG(t)^{-1}g(t) = DG(t)^{-1}F^{*}(\hat{x})$$
$$QG(t)^{-1}\hat{d}^{\circ}(t) = Q\hat{x}(t) + QN(t)D\hat{x}(t) - QG(t)^{-1}g(t) = QG(t)^{-1}F^{*}(\hat{x})$$

Note that $QG(t)^{-1}\hat{d}^{\circ}(t_{i,j}) = 0\,\forall\, i = 1, \ldots, N,\ j = 0, \ldots, s$ owing to collocation conditions and the continuity conditions $\hat{x}(t_{i,0}) = \hat{x}(t_{i-1,s})$.

Analysis of deviation
Differential component

Combining the above equations we obtain for $r := \hat{\epsilon} - \hat{e}$:

$$\frac{Dr_{i,j} - Dr_{i,j-1}}{\delta_{i,j}} = -DN(t_{i,j})Dr_{i,j}$$

$$\underbrace{-DN(t_{i,j})D\hat{e}(t_{i,j}) + \sum_{k=0}^{s} \alpha_{jk} DN(t_{i,k})D\hat{e}(t_{i,k})}_{\leq C\delta_{i,j}(\|D\hat{e}\|+\|D\hat{e}'\|)=\mathcal{O}(h^{s+1})} + \mathcal{O}(h^{s+1})$$

$$+ \sum_{k=0}^{s} \alpha_{jk} \underbrace{\left(DG(t_{i,j})^{-1} - DG(t_{i,k})^{-1}\right)}_{\mathcal{O}(h)} \underbrace{\hat{d}^{\circ}(t_{i,k})}_{\mathcal{O}(h^s)}$$

The claimed order $Dr = \mathcal{O}(h^{s+1})$ now follows from the stability of the backward Euler scheme.

Analysis of deviation
Algebraic component

For the algebraic component we have

$$
\begin{aligned}
Qr_{i,j} &= -QN(t_{i,j})Dr_{i,j} + QG(t_{i,j})^{-1}\hat{d}_{i,j} \\
&= -QN(t_{i,j})Dr_{i,j} + \sum_{k=0}^{s} \alpha_{jk} \underbrace{\left(QG(t_{i,j})^{-1} - QG(t_{i,k})^{-1}\right)}_{\mathcal{O}(h)} \underbrace{\hat{d}^{\circ}(t_{i,k})}_{\mathcal{O}(h^s)} + \\
&\qquad + \sum_{k=0}^{s} \alpha_{jk} \underbrace{QG(t_{i,k})^{-1}\hat{d}^{\circ}(t_{i,k})}_{=0}
\end{aligned}
$$

$QG(t_{i,j})^{-1}\hat{d}^{\circ}(t_{i,j}) = 0 \,\forall\, i = 1, \ldots, N,\ j = 0, \ldots, s$ owing to the collocation and continuity conditions $\hat{x}(t_{i,0}) = \hat{x}(t_{i-1,s})$.

Conclusions and ongoing research

Conclusions

Conclusion

- Defect correction provides a basis for inexpensive and reliable error estimate for DAEs
- Interplay between working schemes and defect evaluation is essential
- In some particular cases asymptotic correctness established
- In other cases numerical evidence suggests asymptotic correctness or acceptable approximation

Conclusions and ongoing research

Open questions

Ongoing research

- Index 2 or higher index problems
- Essential singularities (theory for collocation not yet available)
- Realization in a nonlinear setting, combination with Newton method
- etc.

Quadrature-defect-based a-posteriori error estimates for differential-algebraic equations

W. Auzinger, H. Lehner, E. Weinmüller

Institute for Analysis and Scientific Computing
Vienna University of Technology

Austrian Numerical Analysis Day Linz 2008