Universidad de La Laguna Departamento de Análisis Matemático

On the variable stepsize performance of SAFERK methods.

S. González-Pinto, D. Hernández-Abreu* (Univ. La Laguna, Spain)

> and B. Simeon (*TU Kaiserslautern*, Germany).



ULL

SciCADE 2011 University of Toronto

July 12th, 2011.







OUTLINE

- SAFERK methods: algebraic order and linear stability.
- Convergence on non-stiff problems and stiff semilinear systems.
- Convergence on index one/two DAEs.
- Implementation issues.
- Numerical illustrations.
- Concluding remarks and acknowledgements.





A new family of collocation Runge-Kutta methods

$$\begin{cases} Y_{n,i} &= y_n + h_n \sum_{j=1}^s a_{ij} f(t_n + c_j h_n, Y_{n,j}), & 1 \le i \le s, \\ y_{n+1} &= y_n + h_n \sum_{j=1}^s b_j f(t_n + c_j h_n, Y_{n,j}), \end{cases}$$

with good stability and convergence properties has been recently introduced in

An efficient family of strongly A-stable Runge-Kutta collocation methods for stiff systems and DAEs.

- Part I: Stability and order results. JCAM 2010.
- Part II: Convergence results. To appear in APNUM.



A new family of collocation Runge-Kutta methods

$$\begin{cases} Y_{n,i} &= y_n + h_n \sum_{j=1}^s a_{ij} f(t_n + c_j h_n, Y_{n,j}), & 1 \le i \le s, \\ y_{n+1} &= y_n + h_n \sum_{j=1}^s b_j f(t_n + c_j h_n, Y_{n,j}), \end{cases}$$

with good stability and convergence properties has been recently introduced in

An efficient family of strongly A-stable Runge-Kutta collocation methods for stiff systems and DAEs.

• Part I: Stability and order results. JCAM 2010.

Part II: Convergence results. To appear in APNUM.



The so-called SAFERK methods are competitive regarding RadaullA methods with the same number of implicit stages for stiff systems and index one/two DAEs.

For $s \ge 3$, a uniparametric family of s-stage RK methods is obtained. For each parameter value α , the corresponding $SAFERK(\alpha, s)$ method possesses the following features:



For $s \ge 3$, a uniparametric family of s-stage RK methods is obtained. For each parameter value α , the corresponding $SAFERK(\alpha, s)$ method possesses the following features:

• it is *Stiffly Accurate* method with a *First Explicit* stage:

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array} \qquad \begin{array}{c|c} e_1^T \cdot A = 0^T \Rightarrow Y_{n,1} = y_n \\ e_s^T \cdot A = b^T \Rightarrow Y_{n,s} = y_{n+1} \end{array}$$



For $s \ge 3$, a uniparametric family of s-stage RK methods is obtained. For each parameter value α , the corresponding $SAFERK(\alpha, s)$ method possesses the following features:

• it is *Stiffly Accurate* method with a *First Explicit* stage:

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array} \qquad \begin{array}{c|c} e_1^T \cdot A = 0^T \Rightarrow Y_{n,1} = y_n \\ e_s^T \cdot A = b^T \Rightarrow Y_{n,s} = y_{n+1} \end{array}$$

• it has stage order q = s, i.e, it is a collocation method based on a certain interpolatory quadrature $\{b_i, c_i\}_{i=1}^s$, with $c_1 = 0$ and $c_s = 1$;



For $s \ge 3$, a uniparametric family of *s*-stage RK methods is obtained. For each parameter value α , the corresponding $SAFERK(\alpha, s)$ method possesses the following features:

• it is *Stiffly Accurate* method with a *First Explicit* stage:

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array} \qquad \begin{array}{c|c} e_1^T \cdot A = 0^T \Rightarrow Y_{n,1} = y_n \\ e_s^T \cdot A = b^T \Rightarrow Y_{n,s} = y_{n+1} \end{array}$$

• it has stage order q = s, i.e, it is a collocation method based on a certain interpolatory quadrature $\{b_i, c_i\}_{i=1}^s$, with $c_1 = 0$ and $c_s = 1$;

 $\sqrt{2s+1}(P_s^*(x) - P_{s-2}^*(x)) + \alpha \sqrt{2s-1}(P_{s-1}^*(x) - P_{s-3}^*(x)) = 0$ ($P_n^*(x)$ normalized Legendre polynomials on [0,1], $P_n^*(1) = 1$).



For $s \geq 3$, a uniparametric family of s-stage RK methods is obtained. For each parameter value α , the corresponding $SAFERK(\alpha, s)$ method possesses the following features:

• it is *Stiffly Accurate* method with a *First Explicit* stage:

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array} \qquad \begin{array}{c|c} e_1^T \cdot A = 0^T \Rightarrow Y_{n,1} = y_n \\ \hline & e_s^T \cdot A = b^T \Rightarrow Y_{n,s} = y_{n+1} \end{array}$$

• it has stage order q = s, i.e, it is a collocation method based on a certain interpolatory quadrature $\{b_i, c_i\}_{i=1}^s$, with $c_1 = 0$ and $c_s = 1$;

- Q it has algebraic order p = 2s 3, for all $\alpha \neq 0$;
- it is computationally equivalent to the (s-1)-stage RadaullA method (similar implicitness over each integration step).



For each $s \geq 3$, the linear stability function

$$R(z) = 1 + zb^{T}(I - zA)^{-1}e, \quad e = (1, \dots, 1)^{T} \in \mathbb{R}^{s},$$

of a $SAFERK(\alpha, s)$ method fulfils:



For each $s \geq 3$, the linear stability function

$$R(z) = 1 + zb^{T}(I - zA)^{-1}e, \quad e = (1, \dots, 1)^{T} \in \mathbb{R}^{s},$$

of a $SAFERK(\alpha, s)$ method fulfils:

- $\begin{array}{l} \blacksquare & |R(z)| \leq 1, \ \forall z \in \mathbb{C}^- \ \mbox{(i.e, A-acceptability) if only if $\alpha \leq 0$ and $\alpha \neq -\gamma_s$;} \end{array}$
- A-acceptability $+ |R(\infty)| < 1$ if only if $\alpha < 0$ and $\alpha \neq -\gamma_s$. $R(\infty) = (-1)^{s+1} \frac{\gamma_s + \alpha}{\gamma_s - \alpha}.$



For each $s \geq 3$, the linear stability function

$$R(z) = 1 + zb^{T}(I - zA)^{-1}e, \quad e = (1, \dots, 1)^{T} \in \mathbb{R}^{s},$$

of a $SAFERK(\alpha, s)$ method fulfils:

- $|R(z)| \leq 1, \ \forall z \in \mathbb{C}^-$ (i.e, A-acceptability) if only if $\alpha \leq 0$ and $\alpha \neq -\gamma_s$;
- A-acceptability + $|R(\infty)| < 1$ if only if $\alpha < 0$ and $\alpha \neq -\gamma_s$. $R(\infty) = (-1)^{s+1} \frac{\gamma_s + \alpha}{\gamma_s - \alpha}.$

• Although $SAFERK(\alpha, s)$ are strongly A-stable iff $\alpha < 0$, $\alpha \neq -\gamma_s$, there are not L-stable methods $(|R(\infty)| = 0)$:

$$|R(\infty)| \in \left[\frac{1}{s-1}, 1\right)$$
 for $-\frac{s-2}{s}\gamma_s \le \alpha < 0$.





• The principal term of local error for a Runge-Kutta method $y_{RK}(t+h;t,y(t)) - y(t) = PTLE(t,h) + O(h^{p+2})$

$$PTLE(t,h) = \frac{h^{p+1}}{(p+1)!} \sum_{\tau \in LT_{p+1}} (1 - \omega(\tau)) F(\tau)(y(t))$$



• The principal term of local error for a Runge-Kutta method $y_{RK}(t+h;t,y(t)) - y(t) = PTLE(t,h) + O(h^{p+2})$

$$PTLE(t,h) = \frac{h^{p+1}}{(p+1)!} \sum_{\tau \in LT_{p+1}} (1 - \omega(\tau)) F(\tau)(y(t))$$

• For non-stiff problems, those methods with a smaller l_2 -norm of the error coefficients are preferred:

$$EC_p(RK) := \frac{1}{(p+1)!} \sqrt{\sum_{\tau \in LT_{p+1}} (1 - \omega(\tau))^2}.$$



$$\mathcal{K}_{s}(\alpha) := \frac{EC_{2s-3}(SAFERK(\alpha, s))}{EC_{2s-3}(\text{RadauIIA}(s-1))} < 1.$$

• Theorem. For an s-stage Runge–Kutta method fulfilling B(p), C(q) and D(r), with $p \le \min\{q + r, 2q + 1\}$ we have

 $1 - \omega(\tau) = K(\tau; q, r)(1 - (p+1)b^T c^p), \quad \forall \tau \in LT_{p+1}.$



• Theorem. For an s-stage Runge–Kutta method fulfilling B(p), C(q) and D(r), with $p \leq \min\{q + r, 2q + 1\}$ we have

$$1 - \omega(\tau) = K(\tau; q, r)(1 - (p+1)b^T c^p), \quad \forall \tau \in LT_{p+1}.$$

• For $SAFERK(\alpha, s)$ and RadauIIA(s-1) methods, we have

$$\mathcal{K}_{s}(\alpha) = |\alpha| \left(\frac{s-2}{s}\gamma_{s}\right)^{-1} \sqrt{\frac{\sum_{\tau \in LT_{2s-2}} K(\tau; s, s-3)^{2}}{\sum_{\tau \in LT_{2s-2}} K(\tau; s-1, s-2)^{2}}},$$



• Theorem. For an s-stage Runge–Kutta method fulfilling B(p), C(q) and D(r), with $p \le \min\{q + r, 2q + 1\}$ we have

$$1 - \omega(\tau) = K(\tau; q, r)(1 - (p+1)b^T c^p), \quad \forall \tau \in LT_{p+1}.$$

• For $SAFERK(\alpha, s)$ and RadauIIA(s-1) methods, we have

$$\mathcal{K}_{s}(\alpha) = |\alpha| \left(\frac{s-2}{s}\gamma_{s}\right)^{-1} \sqrt{\frac{\sum_{\tau \in LT_{2s-2}} K(\tau; s, s-3)^{2}}{\sum_{\tau \in LT_{2s-2}} K(\tau; s-1, s-2)^{2}}},$$



• Theorem. For $s \ge 3$, strongly A-stable SAFERK methods with ratio $\mathcal{K}_s(\alpha) < 1$ are obtained if and only if $\alpha_s < \alpha < 0$



• Theorem. For an s-stage Runge–Kutta method fulfilling B(p), C(q) and D(r), with $p \le \min\{q + r, 2q + 1\}$ we have

$$1 - \omega(\tau) = K(\tau; q, r)(1 - (p+1)b^T c^p), \quad \forall \tau \in LT_{p+1}.$$

• For $SAFERK(\alpha, s)$ and RadauIIA(s-1) methods, we have

$$\mathcal{K}_{s}(\alpha) = |\alpha| \left(\frac{s-2}{s}\gamma_{s}\right)^{-1} \sqrt{\frac{\sum_{\tau \in LT_{2s-2}} K(\tau; s, s-3)^{2}}{\sum_{\tau \in LT_{2s-2}} K(\tau; s-1, s-2)^{2}}},$$



• Theorem. For $s \ge 3$, strongly A-stable SAFERK methods with ratio $\mathcal{K}_s(\alpha) < 1$ are obtained if and only if $\alpha_s < \alpha < 0$, where α_s fulfils $\mathcal{K}_s(\alpha_s) = 1$ and $-\gamma_s < \alpha_s < -\frac{s-2}{s}\gamma_s$.

• Theorem. For an s-stage Runge–Kutta method fulfilling B(p), C(q) and D(r), with $p \le \min\{q + r, 2q + 1\}$ we have

$$1 - \omega(\tau) = K(\tau; q, r)(1 - (p+1)b^T c^p), \quad \forall \tau \in LT_{p+1}.$$

• For $SAFERK(\alpha, s)$ and RadauIIA(s-1) methods, we have

$$\mathcal{K}_{s}(\alpha) = |\alpha| \left(\frac{s-2}{s}\gamma_{s}\right)^{-1} \sqrt{\frac{\sum_{\tau \in LT_{2s-2}} K(\tau; s, s-3)^{2}}{\sum_{\tau \in LT_{2s-2}} K(\tau; s-1, s-2)^{2}}},$$

• Theorem. For $s \ge 3$, strongly A-stable SAFERK methods with ratio $\mathcal{K}_s(\alpha) < 1$ are obtained if and only if $\alpha_s < \alpha < 0$, where α_s fulfils $\mathcal{K}_s(\alpha_s) = 1$ and $-\gamma_s < \alpha_s < -\frac{s-2}{s}\gamma_s$.



• Hence, strongly A-stable $SAFERK(\alpha, s)$ methods, with $c_i \in [0, 1]$ and $b_i > 0$ $(1 \le i \le s)$ fulfil $\mathcal{K}_s(\alpha) < 1$.





$$y'(t) = Jy + g(t, y), \quad \mu[J] = \mathcal{O}(1), \ L_g(y) = \mathcal{O}(1).$$



$$y'(t) = Jy + g(t, y), \quad \mu[J] = \mathcal{O}(1), \ L_g(y) = \mathcal{O}(1).$$

• The global error $y_n - y(t_n)$ fulfils

LobattoIIIA(s)	RadauIIA(s-1)	$SAFERK(\alpha, s)$
$\mathcal{O}(h^{\max\{4,s\}})$	$\mathcal{O}(h^s)$	$\mathcal{O}(h^{\min\{s+1,2s-3\}})$



$$y'(t) = Jy + g(t, y), \quad \mu[J] = \mathcal{O}(1), \ L_g(y) = \mathcal{O}(1).$$

• The global error $y_n - y(t_n)$ fulfils

LobattoIIIA(s)	RadauIIA(s-1)	$SAFERK(\alpha, s)$
$\mathcal{O}(h^{\max\{4,s\}})$	$\mathcal{O}(h^s)$	$\mathcal{O}(h^{\min\{s+1,2s-3\}})$



• Optimal B-convergence $\mathcal{O}(h^{q+1})$: for A-, AS- and ASI- stable methods fulfilling B(q+1) and C(q), such that $\psi(z)$ remains uniformly bounded on \mathbb{C}^- .



$$y'(t) = Jy + g(t, y), \quad \mu[J] = \mathcal{O}(1), \ L_g(y) = \mathcal{O}(1).$$

• The global error $y_n - y(t_n)$ fulfils

LobattoIIIA(s)	RadauIIA(s-1)	$SAFERK(\alpha, s)$
$\mathcal{O}(h^{\max\{4,s\}})$	$\mathcal{O}(h^s)$	$\mathcal{O}(h^{\min\{s+1,2s-3\}})$

- E.
- Optimal B-convergence $\mathcal{O}(h^{q+1})$: for A-, AS- and ASI- stable methods fulfilling B(q+1) and C(q), such that $\psi(z)$ remains uniformly bounded on \mathbb{C}^- .

$$\psi(z) = \frac{b^T (I - zA)^{-1} \zeta_q}{b^T (I - zA)^{-1} e}$$
, with $\zeta_q := \frac{1}{q!} \left(\frac{1}{q+1} c^{q+1} - A c^q \right)$,

• For SAFERK methods we have $\psi(\infty) = \frac{\zeta_s^{(s)}}{1-R(\infty)}$.

$$y'(t) = Jy + g(t, y), \quad \mu[J] = \mathcal{O}(1), \ L_g(y) = \mathcal{O}(1).$$

• The global error $y_n - y(t_n)$ fulfils

LobattoIIIA(s)	RadauIIA(s-1)	$SAFERK(\alpha, s)$
$\mathcal{O}(h^{\max\{4,s\}})$	$\mathcal{O}(h^s)$	$\mathcal{O}(h^{\min\{s+1,2s-3\}})$

- Ē
- Optimal B-convergence $\mathcal{O}(h^{q+1})$: for A-, AS- and ASI- stable methods fulfilling B(q+1) and C(q), such that $\psi(z)$ remains uniformly bounded on \mathbb{C}^- .



• ASI-stability (resp. AS-stability): I - zA is regular, $z \in \mathbb{C}^-$, and $(I - zA)^{-1}$ (resp. $zb^T(I - zA)^{-1}$) is uniformly bounded on \mathbb{C}^- .



Convergence on DAEs

1. Index 1 DAEs

 $y'(t) = f(y, z), \ 0 = g(y, z), \ \det(g_z(y, z)) \neq 0.$



FII

1. Index 1 DAEs

 $y'(t) = f(y, z), \ 0 = g(y, z), \ \det(g_z(y, z)) \neq 0.$

• For consistent initial values (y_0, z_0) , i.e. $g(y_0, z_0) = 0$, the advancing solution provided by a Runge-Kutta method fulfils

$$Y_{n,i} = y_n + h \sum_{j=1}^s a_{ij} f(Y_{nj}, Z_{nj}), \quad \mathbf{0} = g(Y_{n,i}, Z_{n,i}), \quad 1 \le i \le s,$$

with $Y_{n,1} = y_n$, $Z_{n,1} = z_n$, $y_{n+1} = Y_{n,s}$ and $z_{n+1} = Z_{n,s}$.



1. Index 1 DAEs

 $y'(t) = f(y, z), \ 0 = g(y, z), \ \det(g_z(y, z)) \neq 0.$

For consistent initial values (y_0, z_0) , i.e. $g(y_0, z_0) = 0$, the advancing solution provided by a Runge-Kutta method fulfils

$$Y_{n,i} = y_n + h \sum_{j=1}^s a_{ij} f(Y_{nj}, Z_{nj}), \quad \mathbf{0} = g(Y_{n,i}, Z_{n,i}), \quad 1 \le i \le s,$$

with
$$Y_{n,1} = y_n$$
, $Z_{n,1} = z_n$, $y_{n+1} = Y_{n,s}$ and $z_{n+1} = Z_{n,s}$.

• For stiffly accurate methods, the numerical solutions are equivalent to those obtained from the ODE y' = f(y, G(y)), with z = G(y).

1. Index 1 DAEs

 $y'(t) = f(y, z), \ 0 = g(y, z), \ \det(g_z(y, z)) \neq 0.$

For consistent initial values (y_0, z_0) , i.e. $g(y_0, z_0) = 0$, the advancing solution provided by a Runge-Kutta method fulfils

$$Y_{n,i} = y_n + h \sum_{j=1}^s a_{ij} f(Y_{nj}, Z_{nj}), \quad \mathbf{0} = g(Y_{n,i}, Z_{n,i}), \quad 1 \le i \le s,$$

with $Y_{n,1} = y_n$, $Z_{n,1} = z_n$, $y_{n+1} = Y_{n,s}$ and $z_{n+1} = Z_{n,s}$.

 \bigcirc For both components y and z we have full order p

LobattoIIIA(s)RadauIIA(s-1)SAFERK(
$$\alpha, s$$
) $\mathcal{O}(h^{2s-2})$ $\mathcal{O}(h^{2s-3})$ $\mathcal{O}(h^{2s-3})$

Convergence on DAEs

2. Index 2 DAEs

 $y'(t) = f(y, z), \ 0 = g(y), \ \det(g_y \cdot f_z)(y, z) \neq 0.$



FI

Convergence on DAEs

2. Index 2 DAEs

 $y'(t) = f(y, z), \ 0 = g(y), \ \det(g_y \cdot f_z)(y, z) \neq 0.$

Q. Since $e_1^T A = 0^T$ and the submatrix \tilde{A} is regular

$$Y_{ni} = y_n + h \sum_{j=1}^{s} a_{ij} f(Y_{nj}, Z_{nj}), \quad 0 = g(Y_{ni}), \qquad 1 \le i \le s$$

admits a locally unique solution $\{(Y_{ni}, Z_{ni})\}_{i=1}^{s}$ such that $Y_{n1} = y_n$ and $Z_{n1} = z_n$.

Q. We have that $z_{n+1} = Z_{ns}$ and $y_{n+1} = Y_{ns}$, with $g(y_{n+1}) = 0$.
2. Index 2 DAEs

 $y'(t) = f(y, z), \ 0 = g(y), \ \det(g_y \cdot f_z)(y, z) \neq 0.$

- L. Jay (BIT, 1993): global error estimates for the whole family of stiffly accurate methods with a first internal stage of explicit type and a regular submatrix $\tilde{A} = (a_{ij})_{2 \le i,j \le s}$.
- For consistent initial values (y_0, z_0) , we get

LobattoIIIA(s)	RadauIIA(s-1)	$SAFERK(\alpha, s)$
$\mathcal{O}(h^{2s-2}, \frac{h^{s-1}}{n})$	$\mathcal{O}(h^{2s-3}, h^{s-1})$	$\mathcal{O}(h^{2s-3}, \mathbf{h^s})$







1. Solving the stage equation:



1. Solving the stage equation:

Since
$$A = \begin{pmatrix} 0 & 0^T \\ A_1 & \tilde{A} \end{pmatrix}$$
, the stage equation
 $Y = e \otimes y_n + h(A \otimes I)F(t_n, Y)$,
with $Y = (Y_1^T, \dots, Y_4^T)^T$, $Y_i \approx y(t_n + hc_i)$, reduces to
 $Z - h(A_1 \otimes I)f(t_n, y_n) - h(\tilde{A} \otimes I)F(Z) = 0$
with $Z = (Z_2^T, Z_3^T, Z_4^T)^T$, and $Z_i = Y_i - y_n$.

4



1. Solving the stage equation:

Since
$$A = \begin{pmatrix} 0 & 0^T \\ A_1 & \tilde{A} \end{pmatrix}$$
, the stage equation
 $Y = e \otimes y_n + h(A \otimes I)F(t_n, Y)$,
with $Y = (Y_1^T, \dots, Y_4^T)^T$, $Y_i \approx y(t_n + hc_i)$, reduces to
 $Z - h(A_1 \otimes I)f(t_n, y_n) - h(\tilde{A} \otimes I)F(Z) = 0$
with $Z = (Z_2^T, Z_3^T, Z_4^T)^T$, and $Z_i = Y_i - y_n$.

The simplified Newton iteration then reads as

$$(I - h(\tilde{A} \otimes J))\Delta Z^{(\nu)} = -Z^{(\nu)} + h(A_1 \otimes I)f(t_n, y_n) + h(\tilde{A} \otimes I)F(Z^{(\nu)})$$
$$J = \frac{\partial f}{\partial y}(t_n, y_n), \quad \Delta Z^{(\nu)} = Z^{(\nu+1)} - Z^{(\nu)}.$$

Since
$$T^{-1}\tilde{A}^{-1}T = \Lambda = \begin{pmatrix} \gamma & 0 & 0 \\ 0 & \delta & -\omega \\ 0 & \omega & \delta \end{pmatrix}$$
, the iteration process is

$$(h^{-1}\Lambda \otimes I - I \otimes J)\Delta W^{(\nu)} = -h^{-1}(\Lambda \otimes I)W^{(\nu)} + \Lambda T^{-1}A_1 \otimes f(t_n, y_n) + (T^{-1} \otimes I)F((T \otimes I)W^{(\nu)})$$

with $W^{(\nu)} := (T^{-1} \otimes I) Z^{(\nu)}$ and $\Delta W^{(\nu)} = W^{(\nu+1)} - W^{(\nu)}$.



Since
$$T^{-1}\tilde{A}^{-1}T = \Lambda = \begin{pmatrix} \gamma & 0 & 0 \\ 0 & \delta & -\omega \\ 0 & \omega & \delta \end{pmatrix}$$
, the iteration process is

$$(h^{-1}\Lambda \otimes I - I \otimes J)\Delta W^{(\nu)} = -h^{-1}(\Lambda \otimes I)W^{(\nu)} + \Lambda T^{-1}A_1 \otimes f(t_n, y_n) + (T^{-1} \otimes I)F((T \otimes I)W^{(\nu)})$$

with $W^{(\nu)} := (T^{-1} \otimes I)Z^{(\nu)}$ and $\Delta W^{(\nu)} = W^{(\nu+1)} - W^{(\nu)}$.

• In particular, this linear system requires a LU-decomposition for the matrix $(h^{-1}\gamma I - J)$ at each integration step.



• The corresponding iteration for the RadaullA method is essentially the same as for SAFERK methods but with $A_1 = 0$.

$$(h^{-1}\Lambda \otimes I - I \otimes J)\Delta W^{(\nu)} = -h^{-1}(\Lambda \otimes I)W^{(\nu)} + \Lambda T^{-1}A_1 \otimes f(t_n, y_n) + (T^{-1} \otimes I)F((T \otimes I)W^{(\nu)}).$$

• The iterative scheme is stopped at the first iteration r, such that

$$\max_{2 \le i \le 4} \left\| Z_i^r - Z_i^{r-1} \right\| \le c \cdot Tol, \ c := 0'03.$$



$$(h^{-1}\Lambda \otimes I - I \otimes J)\Delta W^{(\nu)} = -h^{-1}(\Lambda \otimes I)W^{(\nu)} + \Lambda T^{-1}A_1 \otimes f(t_n, y_n) + (T^{-1} \otimes I)F((T \otimes I)W^{(\nu)}).$$



$$\max_{2 \le i \le 4} \left\| Z_i^r - Z_i^{r-1} \right\| \le c \cdot Tol, \ c := 0'03.$$

• For the first integration step n = 0 and the first iterate $\nu = 0$ we consider $Z^{(0)} = 0$ (i.e., $W^{(0)} = 0$) and then

 $(h^{-1}\Lambda \otimes I - I \otimes J)W^{(1)} = (\Lambda T^{-1}A_1 + T^{-1}\tilde{e}) \otimes f(t_0, y_0)$



$$(h^{-1}\Lambda \otimes I - I \otimes J)\Delta W^{(\nu)} = -h^{-1}(\Lambda \otimes I)W^{(\nu)} + \Lambda T^{-1}A_1 \otimes f(t_n, y_n) + (T^{-1} \otimes I)F((T \otimes I)W^{(\nu)}).$$



$$\max_{2 \le i \le 4} \left\| Z_i^r - Z_i^{r-1} \right\| \le c \cdot Tol, \ c := 0'03.$$



$$Z_{i,n+1}^{(0)} = q(t_{n+1} + c_i h_{n+1}) + y_n - y_{n+1}, \ 1 \le i \le 4,$$

with $q(t) \in \Pi_3$ such that $q(t_n) = 0$ and $q(t_n + c_i h_n) = Z_{i,n}$, $2 \le i \le 4$.

2. Embedded formula for the local error estimation:

2. Embedded formula for the local error estimation:

• A fourth order formula cannot be embedded to a given $SAFERK(\alpha, 4)$ method (of order 5).

2. Embedded formula for the local error estimation:

• For each $SAFERK(\alpha, 4)$ method, a one-parameter family of third order methods can be embedded.

0	0	0	0	0
c_2	a_{21}		•	
c_3	a_{31}		Ã	
c_4	a_{41}			
(5)	b_1	b_2	b_3	b_4
(3)	d_1	d_2	d_3	d_4



2. Embedded formula for the local error estimation:

• For each $SAFERK(\alpha, 4)$ method, a one-parameter family of third order methods can be embedded.

0	0	0	0	0
c_2	a_{21}			
c_3	a_{31}		Ã	
c_4	a_{41}			
(5)	b_1	b_2	b_3	b_4
(3)	d_1	d_2	d_3	d_4



The local error estimation for the RadauIIA(3) requires an extra function evaluation at each integration point f(t_n, y_n) (by adding a first stage of explicit type).

2. Embedded formula for the local error estimation:

• For each $SAFERK(\alpha, 4)$ method, a one-parameter family of third order methods can be embedded.

0	0	0	0	0
c_2	a_{21}			
c_3	a_{31}		Ã	
c_4	a_{41}			
(5)	b_1	b_2	b_3	b_4
(3)	d_1	d_2	d_3	d_4



• Hence, the local error estimation for both $SAFERK(\alpha, 4)$ and RadauIIA(3) methods requires the same number of function evaluations at each integration step.

From the stage equation for the $SAFERK(\alpha, 4)$ method:

$$hF(Z) = (\tilde{A}^{-1} \otimes I) (Z - hA_1 \otimes f(t_n, y_n)).$$



From the stage equation for the $SAFERK(\alpha, 4)$ method:

$$hF(Z) = (\tilde{A}^{-1} \otimes I) (Z - hA_1 \otimes f(t_n, y_n)).$$

The local error estimator

 $\hat{y}_{n+1} - y_{n+1} = h(\mathbf{d}_1 - b_1)f(t_n, y_n) + h\sum_{i=2}^4 (d_i - b_i)f(t_n + c_i h, y_n + Z_i)$

can be expressed as
$$\hat{y}_{n+1} - y_{n+1} = hf(t_n, y_n) \cdot e_1 + \sum_{i=2}^4 e_i \cdot Z_i$$
,

with

$$e_1 := (d_1 - b_1) - (\tilde{d} - \tilde{b})^T \tilde{A}^{-1} A_1,$$

$$(e_2, e_3, e_4)^T := (\tilde{d} - \tilde{b})^T \tilde{A}^{-1}.$$



However, on linear problems $y'=\lambda y$, this local error estimator is unbounded for $z=h\lambda\to\infty$

$$\hat{y}_{n+1} - y_{n+1} \approx e_1 z y_n.$$

– p. 23/43

However, on linear problems $y' = \lambda y$, this local error estimator is unbounded for $z = h\lambda \to \infty$

$$\hat{y}_{n+1} - y_{n+1} \approx e_1 z y_n.$$

A filtering process is then considered

$$err = (I - h\gamma^{-1}J)^{-1}(\hat{y}_{n+1} - y_{n+1})$$

Recall that an LU-decomposition for $(h^{-1}\gamma I - J)$ is already available.



However, on linear problems $y' = \lambda y$, this local error estimator is unbounded for $z = h\lambda \to \infty$

$$\hat{y}_{n+1} - y_{n+1} \approx e_1 z y_n.$$

A filtering process is then considered

$$err = (I - h\gamma^{-1}J)^{-1}(\hat{y}_{n+1} - y_{n+1})$$

Recall that an LU-decomposition for $(h^{-1}\gamma I - J)$ is already available.

• Observe that we still have $err = \mathcal{O}(h^4)$ for $h \to 0$, whereas for linear problems and $z \to \infty$

$$err \to -(e_1\gamma)y_n.$$



• A second filtering is done after rejections with ||err|| > 1:

$$\widetilde{err} = (I - h\gamma^{-1}J)^{-1} \left(e_1 hf(t_n, y_n + err) + \sum_{i=2}^4 e_i Z_i \right)$$

– p. 24/43

Q A second filtering is done after rejections with ||err|| > 1:

$$\widetilde{err} = (I - h\gamma^{-1}J)^{-1} \left(e_1 hf(t_n, y_n + err) + \sum_{i=2}^4 e_i Z_i \right)$$

Q. Now, for linear problems and $z \to \infty$

$$\widetilde{err} = \frac{1 - z(\gamma^{-1} - e_1)}{1 - z\gamma^{-1}} err \to 0 \iff e_1 = \gamma^{-1}.$$



• A second filtering is done after rejections with ||err|| > 1:

$$\widetilde{err} = (I - h\gamma^{-1}J)^{-1} \left(e_1 hf(t_n, y_n + err) + \sum_{i=2}^4 e_i Z_i \right)$$

Q. Now, for linear problems and $z \to \infty$

$$\widetilde{err} = \frac{1 - z(\gamma^{-1} - e_1)}{1 - z\gamma^{-1}} err \to 0 \iff e_1 = \gamma^{-1}.$$

C This latter condition determines a unique embedded method (i.e., the parameter d_1) for the underlying $SAFERK(\alpha, 4)$ method.



• A second filtering is done after rejections with ||err|| > 1:

$$\widetilde{err} = (I - h\gamma^{-1}J)^{-1} \left(e_1 hf(t_n, y_n + err) + \sum_{i=2}^4 e_i Z_i \right)$$

Q. Now, for linear problems and $z \to \infty$

$$\widetilde{err} = \frac{1 - z(\gamma^{-1} - e_1)}{1 - z\gamma^{-1}} err \to 0 \iff e_1 = \gamma^{-1}.$$

The filtered local error estimator is then computed from

$$(h^{-1}\gamma I - J)err = f(t_n, y_n) + h^{-1}\sum_{i=2}^{4} (e_i\gamma)Z_i.$$



• A second filtering is done after rejections with ||err|| > 1:

$$\widetilde{err} = (I - h\gamma^{-1}J)^{-1} \left(e_1 hf(t_n, y_n + err) + \sum_{i=2}^4 e_i Z_i \right)$$

Q. Now, for linear problems and $z \to \infty$

$$\widetilde{err} = \frac{1 - z(\gamma^{-1} - e_1)}{1 - z\gamma^{-1}} err \to 0 \iff e_1 = \gamma^{-1}.$$

The filtered local error estimator is then computed from

$$(h^{-1}\gamma I - J)err = f(t_n, y_n) + h^{-1}\sum_{i=2}^{4} (e_i\gamma)Z_i.$$

The stepsize prediction is done under the same conditions as for the RADAU5 code, preferably Gustaffson's controller:

 $\|err_{n+1}\| \approx C_n h_n^4,$

$$\frac{C_{n+1}}{C_n} \approx \frac{C_n}{C_{n-1}}.$$





- We present efficiency plots for the RADAU5 code and a RADAU5-based implementation for some selected 4-stage SAFERK methods on several test problems.
 - Comparisons regarding the variable order code RADAU (based on RadaullA methods of orders 5,9,13) will be also drawn.



- We present efficiency plots for the RADAU5 code and a RADAU5-based implementation for some selected 4-stage SAFERK methods on several test problems.
- Comparisons regarding the variable order code RADAU (based on RadaullA methods of orders 5,9,13) will be also drawn.
- The free parameter α defining the 4-stage SAFERKmethods can be substituted by β , standing for the node c_3 $(0 = c_1 < c_2(\beta) < c_3 := \beta < 1 = c_4)$:

$$R(\infty,\beta) = \frac{(\beta-1)(5\beta-2)}{\beta(5\beta-3)}$$

$$\mathcal{K}_{4}(\beta) := \frac{EC_{5}(SAFERK(\beta, 4))}{EC_{5}(RadauIIA(3))} = 2\sqrt{\frac{58}{103}} \cdot \frac{|1 - 5\beta + 5\beta^{2}|}{|1 - 2\beta|}.$$



$$\beta_1 = \frac{5+\sqrt{5}}{10} \doteq 0'723, \quad \beta_2 = \frac{1}{2} + \frac{1}{10}\sqrt{\frac{248+\sqrt{40479}}{29}} \doteq 0'893.$$

• A-stable SAFERK methods with $\mathcal{K}_4(\beta) \leq 1$ are obtained iff $\beta_1 \leq \beta \leq \beta_2$, with

$$\beta_1 = \frac{5+\sqrt{5}}{10} \doteq 0'723, \quad \beta_2 = \frac{1}{2} + \frac{1}{10}\sqrt{\frac{248+\sqrt{40479}}{29}} \doteq 0'893.$$

Depending on the kind of problem under consideration, it could be preferred to have either more damping at infinity or smaller error coefficients in the principal term of the local error.



• A-stable SAFERK methods with $\mathcal{K}_4(\beta) \leq 1$ are obtained iff $\beta_1 \leq \beta \leq \beta_2$, with

 $\beta_1 = \frac{5+\sqrt{5}}{10} \doteq 0'723, \quad \beta_2 = \frac{1}{2} + \frac{1}{10}\sqrt{\frac{248+\sqrt{40479}}{29}} \doteq 0'893.$

- Depending on the kind of problem under consideration, it could be preferred to have either more damping at infinity or smaller error coefficients in the principal term of the local error.
- For numerical illustrations regarding fixed stepsize integrations, in JCAM2010 we consider

METHOD	β	$R(\infty,eta)$	$\mathcal{K}(eta)$
SAFERK1	0.73	$-0.9388\ldots$	$0.0473\ldots$
SAFERK2	0.74	-0.8532	0.1188
SAFERK3	0.75	$-0.7777\ldots$	$0.1876\ldots$



In order to balance the influence of the damping at infinity and the principal term of local error, we consider the following optimization options:

•
$$\min_{\beta \in [\beta_1, \beta_2]} |R(\infty, \beta)| + \mathcal{K}(\beta) \hookrightarrow \mathsf{SAFERK4};$$

• $\min_{\beta \in [\beta_1, \beta_2]} \max\{|R(\infty, \beta)|, \mathcal{K}(\beta)\} \hookrightarrow \mathsf{SAFERK5}$

In order to balance the influence of the damping at infinity and the principal term of local error, we consider the following optimization options:

•
$$\min_{\beta \in [\beta_1, \beta_2]} |R(\infty, \beta)| + \mathcal{K}(\beta) \hookrightarrow \mathsf{SAFERK4};$$

- $\min_{\beta \in [\beta_1, \beta_2]} \max\{|R(\infty, \beta)|, \mathcal{K}(\beta)\} \hookrightarrow \mathsf{SAFERK5}$
- Summing up:

METHOD	eta	$R(\infty,eta)$	$\mathcal{K}(eta)$
SAFERK1	0.73	-0.9388	0.0473
SAFERK2	0.74	-0.8532	0.1188
SAFERK3	0.75	-0.7777	$0.1876\ldots$
SAFERK4	$0.7566\ldots$	-0.7325	$0.2317\ldots$
SAFERK5	$0.79997\ldots$	-0.5001	$0.5001\ldots$

The Ring Modulator:

The problem comes from electrical circuit analysis and describes the behavior of the ring modulator for a given circuit diagram with 7 capacitors and 8 inductors.

The Ring Modulator:

FII

The problem comes from electrical circuit analysis and describes the behavior of the ring modulator for a given circuit diagram with 7 capacitors and 8 inductors.



The Ring Modulator:

- The problem comes from electrical circuit analysis and describes the behavior of the ring modulator for a given circuit diagram with 7 capacitors and 8 inductors.
- For nonzero values of a certain free-parameter associated to the circuit, the problem reduces to a system of 15 ODEs

$$y' = f(t, y), \quad y \in \mathbb{R}^{15}, \quad t \in [0, 10^{-3}].$$


The Ring Modulator:

- The problem comes from electrical circuit analysis and describes the behavior of the ring modulator for a given circuit diagram with 7 capacitors and 8 inductors.
- For nonzero values of a certain free-parameter associated to the circuit, the problem reduces to a system of 15 ODEs

$$y' = f(t, y), \quad y \in \mathbb{R}^{15}, \quad t \in [0, 10^{-3}].$$

For the work-precision diagrams, we have used:

$$rtol = atol = 10^{-(7+m/4)}, \ 0 \le m \le 20,$$

$$h_0 = 10^{-2} \cdot rtol.$$

The Ring Modulator:

- The problem comes from electrical circuit analysis and describes the behavior of the ring modulator for a given circuit diagram with 7 capacitors and 8 inductors.
- For nonzero values of a certain free-parameter associated to the circuit, the problem reduces to a system of 15 ODEs

$$y' = f(t, y), \quad y \in \mathbb{R}^{15}, \quad t \in [0, 10^{-3}].$$

For the work-precision diagrams, we have used:

$$rtol = atol = 10^{-(7+m/4)}, \ 0 \le m \le 20,$$

$$h_0 = 10^{-2} \cdot rtol.$$



• The code RADAU failed at the tolerances for $0 \le m \le 15$.

The Ring Modulator:



The two Transistor Amplifier:

The problem originates from electrical circuit analysis and it is a model for a two transistor amplifier.

The two Transistor Amplifier:

The problem originates from electrical circuit analysis and it is a model for a two transistor amplifier.



The two Transistor Amplifier:

- The problem originates from electrical circuit analysis and it is a model for a two transistor amplifier.
- Let is a stiff DAE of index 1 and dimension 8

 $My' = f(t, y), \quad y \in \mathbb{R}^8, \quad t \in [0, 0.2].$

where the matrix M is tridiagonal and has rank 5. All components of y are of index 1.



The two Transistor Amplifier:

- The problem originates from electrical circuit analysis and it is a model for a two transistor amplifier.
- Let is a stiff DAE of index 1 and dimension 8

 $My' = f(t, y), \quad y \in \mathbb{R}^8, \quad t \in [0, 0.2].$

where the matrix M is tridiagonal and has rank 5. All components of y are of index 1.

For the work-precision diagrams, we used:

$$rtol = atol = 10^{-(6+m/4)}, \ 0 \le m \le 16,$$

$$h_0 = 10^{-2} \cdot rtol.$$

The two Transistor Amplifier:



The two Transistor Amplifier:



The problem is an index 2 DAE of dimension 49, and models the water flow through a tube system, by considering turbulence and the roughness of the tube walls:

The problem is an index 2 DAE of dimension 49, and models the water flow through a tube system, by considering turbulence and the roughness of the tube walls:



The problem is an index 2 DAE of dimension 49, and models the water flow through a tube system, by considering turbulence and the roughness of the tube walls:

$$My' = f(t, y), \ y \in \mathbb{R}^{49}, \ t \in [0, 17 \cdot 3600], \quad M = \begin{pmatrix} M^{\phi} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & M^p \end{pmatrix}$$



The problem is an index 2 DAE of dimension 49, and models the water flow through a tube system, by considering turbulence and the roughness of the tube walls:

$$My' = f(t, y), \ y \in \mathbb{R}^{49}, \ t \in [0, 17 \cdot 3600], \ M = \begin{pmatrix} M^{\phi} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & M^p \end{pmatrix}$$

where $M^{\phi} \in \mathbb{R}^{18,18}$ is diagonal, and $M^p \in \mathbb{R}^{13,13}$ only has nonzero elements M_{11}^p and M_{22}^p .



The problem is an index 2 DAE of dimension 49, and models the water flow through a tube system, by considering turbulence and the roughness of the tube walls:

$$My' = f(t, y), \ y \in \mathbb{R}^{49}, \ t \in [0, 17 \cdot 3600], \ M = \begin{pmatrix} M^{\phi} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & M^p \end{pmatrix}$$

The first 38 components of y are of index 1, whereas the last 11 are of index 2.



The problem is an index 2 DAE of dimension 49, and models the water flow through a tube system, by considering turbulence and the roughness of the tube walls:

$$My' = f(t,y), \ y \in \mathbb{R}^{49}, \ t \in [0, 17 \cdot 3600], \ M = \begin{pmatrix} M^{\phi} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & M^p \end{pmatrix}$$

- ET.
- The first 38 components of y are of index 1, whereas the last 11 are of index 2.
- The problem has been integrated with

$$rtol = 10^{-(6+m/4)}, \ h_0 = atol = rtol, \ 0 \le m \le 24.$$



RADAU failed for $m = 0, \ldots, 6, 8, 9, 11, \ldots, 14, 16, \ldots, 20, 24$.



The Plate Problem:

 $u_{tt} + \omega u_t + \sigma \Delta \Delta u = f(x, y, t), \quad (x, y) \in \Omega, \ 0 \le t \le 7,$



FII

$$u_{tt} + \omega u_t + \sigma \Delta \Delta u = f(x, y, t), \quad (x, y) \in \Omega, \ 0 \le t \le 7,$$

$$\Omega = \{ (x, y) / \ 0 \le x \le 2, \ 0 \le y \le 4/3 \}$$

$$u|_{\partial\Omega=0}, \ \Delta u|_{\partial\Omega=0}, \ u(x,y,0) = 0, \ u_t(x,y,0) = 0.$$

$$f(x, y, t) = \begin{cases} 200(e^{-5(t-x-2)^2} + e^{-5(t-x-5)^2}), & \text{if } y = 4/9, 8/9, \\ 0, & \text{otherwise.} \end{cases}$$



The Plate Problem:

$$u_{tt} + \omega u_t + \sigma \Delta \Delta u = f(x, y, t), \quad (x, y) \in \Omega, \ 0 \le t \le 7,$$

• We consider the grid $x_i = i\tau$ $(0 \le i \le 9)$, $y_j = j\tau$ $(0 \le j \le 6)$, with $\tau = 2/9$, whereas $\Delta\Delta$ is discretized by means of

The Plate Problem:

 $u_{tt} + \omega u_t + \sigma \Delta \Delta u = f(x, y, t), \quad (x, y) \in \Omega, \ 0 \le t \le 7,$

• An ODE of dimension 80 with Jacobian eigenvalues in the wedge $\{z / Arg(-z) \le 71^{\circ}, -500 \le \text{Re } z < 0\}$ is obtained for $\omega = 1000$ and $\sigma = 100$.



$$u_{tt} + \omega u_t + \sigma \Delta \Delta u = f(x, y, t), \quad (x, y) \in \Omega, \ 0 \le t \le 7,$$

- An ODE of dimension 80 with Jacobian eigenvalues in the wedge $\{z / Arg(-z) \le 71^{\circ}, -500 \le \text{Re } z < 0\}$ is obtained for $\omega = 1000$ and $\sigma = 100$.
- This problem has been integrated in the interval [0,7] with tolerances

$$rtol = 10^{-(2+m/4)}, \ atol = 10^{-3} \cdot rtol, \ 0 \le m \le 32,$$

$$h_0 = 10^{-2} \cdot rtol.$$







$$\begin{cases} y_1'(t) = -0.04y_1(t) + 10^4 y_2 y_3, & y_1(0) = 1, \\ y_2'(t) = 0.04y_1(t) - 10^4 y_2 y_3 - 3 \cdot 10^7 y_2^2, & y_2(0) = 0, \\ y_3'(t) = 3 \cdot 10^7 y_2^2, & y_3(0) = 0. \end{cases}$$



$$y_1'(t) = -0.04y_1(t) + 10^4 y_2 y_3, \qquad y_1(0) = 1,$$

$$y_2'(t) = 0.04y_1(t) - 10^4 y_2 y_3 - 3 \cdot 10^7 y_2^2, \qquad y_2(0) = 0,$$

$$y_3'(t) = 3 \cdot 10^7 y_2^2, \qquad y_3(0) = 0.$$

This problem has an initial transient phase close to t = 0. Moreover, it has a semi-stable equilibrium, which gives rise to unstable integrations in large intervals for non Strongly A-stable methods.



$$y_1'(t) = -0.04y_1(t) + 10^4 y_2 y_3, \qquad y_1(0) = 1,$$

$$y_2'(t) = 0.04y_1(t) - 10^4 y_2 y_3 - 3 \cdot 10^7 y_2^2, \qquad y_2(0) = 0,$$

$$y_3'(t) = 3 \cdot 10^7 y_2^2, \qquad y_3(0) = 0.$$

- This problem has an initial transient phase close to t = 0. Moreover, it has a semi-stable equilibrium, which gives rise to unstable integrations in large intervals for non Strongly A-stable methods.
- It has been integrated in $t \in [0, 10^{11}]$ with

$$rtol = 10^{-(4+m/4)}, \ 0 \le m \le 32,$$

 $atol = 10^{-2} \cdot rtol, \text{ and } atol = 10^{-4} \cdot rtol,$
 $h_0 = 10^{-2} \cdot rtol.$









Adaptive strongly A-stable 4-stage SAFERK methods have been shown to be competitive to the RadauIIA(3) method when implemented in a similar fashion as in the RADAU5 code by E. Hairer and G. Wanner.

Adaptive strongly A-stable 4-stage SAFERK methods have been shown to be competitive to the RadauIIA(3) method when implemented in a similar fashion as in the RADAU5 code by E. Hairer and G. Wanner.

Adaptive SAFERK methods have been tested on 23 problems from the Test Set for IVP Solvers (Univ. Bari, Italy) http://pitagora.dm.uniba.it/ testset/

and the Ernst Hairer's website

http://www.unige.ch/ hairer/testset/testset.html.

SAFERKn	clearly improves	RADAU5	10/23
	slightly improves		8/23
	similar to		3/23
	worse than		2/23

Regarding the variable order code RADAU, adaptive SAFERK methods with enough damping at infinity turn out to be competitive and perform similarly on most of problems when considering medium tolerances.

- Regarding the variable order code RADAU, adaptive SAFERK methods with enough damping at infinity turn out to be competitive and perform similarly on most of problems when considering medium tolerances.
- For stringent tolerances, the RADAU code reflects the combination of higher order (*RADAUIIA*) methods, and it is clearly advantageous over both *SAFERK* and *RADAU5*.





Acknowledgements

We are indebted to Professor E. Hairer for his suggestions and for giving us permission to use and modify the code RADAU5 in order to test the performance of SAFERK methods.



Acknowledgements

We are indebted to Professor E. Hairer for his suggestions and for giving us permission to use and modify the code RADAU5 in order to test the performance of SAFERK methods.

Many thanks for your attention.

