

## 4 Runge-Kutta-Verfahren

### 4.1 Konstruktion

Ausgangspunkt wie immer (Substitution:  $s = t + \tau h$ ,  $0 \leq \tau \leq 1$ )

$$\begin{aligned} \mathbf{y}(t+h) &= \mathbf{y}(t) + [\mathbf{y}(t+h) - \mathbf{y}(t)] = \mathbf{y}(t) + \int_t^{t+h} \mathbf{y}'(s) ds \\ &= \mathbf{y}(t) + h \int_0^1 \mathbf{y}'(t + \tau h) d\tau. \end{aligned}$$

Approximiere durch Quadraturformel

$$\int_0^1 g(\tau) d\tau \approx \sum_{j=1}^m \beta_j g(\gamma_j). \quad (*)$$

Damit zumindest  $g \equiv 1$  exakt integriert wird, fordern wir  $\sum_{j=1}^m \beta_j = 1$ .

Daraus folgt

$$\begin{aligned} \mathbf{y}(t+h) &\approx \mathbf{y}(t) + h \sum_{j=1}^m \beta_j \mathbf{y}'(t + \gamma_j h) \\ &= \mathbf{y}(t) + h \sum_{j=1}^m \beta_j \mathbf{f}(t + \gamma_j h, \mathbf{y}(t + \gamma_j h)). \end{aligned} \quad (\text{RK-1})$$

Problem:  $\mathbf{y}(t + \gamma_j h) = \mathbf{y}(t) + h \int_0^{\gamma_j} \mathbf{y}'(t + \tau h) d\tau$  sind unbekannt. Näherungen wieder durch Quadraturformeln, aber mit den alten Knoten  $\gamma_j$  ( $j = 1, \dots, m$ ) aus (\*) (sonst würden sich neue „Unbekannte“  $\mathbf{y}(t + \text{Knoten} \cdot h)$  ergeben).

$$\int_0^{\gamma_j} g(\tau) d\tau \approx \sum_{\ell=1}^m \alpha_{j,\ell} g(\gamma_\ell) \quad (j = 1, \dots, m). \quad (**)$$

Damit zumindest  $g \equiv 1$  exakt integriert wird, fordern wir  $\sum_{\ell=1}^m \alpha_{j,\ell} = \gamma_j$   
 $\forall j = 1, \dots, m$ .

Damit ergibt sich

$$\begin{aligned} \mathbf{y}(t + \gamma_j h) &\approx \mathbf{y}(t) + h \sum_{\ell=1}^m \alpha_{j,\ell} \mathbf{y}'(t + \gamma_\ell h) \\ &= \mathbf{y}(t) + h \sum_{\ell=1}^m \alpha_{j,\ell} \mathbf{f}(t + \gamma_\ell h, \mathbf{y}(t + \gamma_\ell h)). \end{aligned} \quad (\text{RK-2})$$

Abkürzung:  $\tilde{\mathbf{k}}_j := \mathbf{f}(t + \gamma_j h, \mathbf{y}(t + \gamma_j h))$  ( $j = 1, \dots, m$ ).

(RK-2):  $\tilde{\mathbf{k}}_j \approx \mathbf{f}\left(t + \gamma_j h, \mathbf{y}(t) + h \sum_{\ell=1}^m \alpha_{j,\ell} \tilde{\mathbf{k}}_\ell\right)$  ( $j = 1, \dots, m$ ).

(RK-1):  $\mathbf{y}(t + h) \approx \mathbf{y}(t) + h \sum_{j=1}^m \beta_j \tilde{\mathbf{k}}_j$ .

**$m$ -stufiges Runge-Kutta-Verfahren (RKV):**

$$\begin{aligned} \mathbf{y}_{n+1} &= \mathbf{y}_n + h \sum_{j=1}^m \beta_j \mathbf{k}_j \quad \text{mit} \\ \mathbf{k}_j &= \mathbf{f}\left(t_n + \gamma_j h, \mathbf{y}_n + h \sum_{\ell=1}^m \alpha_{j,\ell} \mathbf{k}_\ell\right) \quad (j = 1, \dots, m). \end{aligned} \quad (\text{RKV})$$

**Butcher-Matrix**

(John Charles Butcher, \*1933)

(auch **Runge-Kutta-ABC**):

$\gamma_1$	$\alpha_{1,1}$	$\cdots$	$\alpha_{1,m}$
$\vdots$	$\vdots$		$\vdots$
$\gamma_m$	$\alpha_{m,1}$	$\cdots$	$\alpha_{m,m}$
	$\beta_1$	$\cdots$	$\beta_m$

**Beispiele.**

0	0	0
1	1	0
	1/2	1/2

symbolisiert ein zweistufiges explizites RKV

(ein RKV ist **explizit**, wenn  $\alpha_{j,\ell} = 0 \ \forall j \leq \ell$  gilt), nämlich das verbesserte Euler-Verfahren.

0	1/4	-1/4	symbolisiert ein zweistufiges implizites RKV:
2/3	1/4	5/12	
	1/4	3/4	

$$\mathbf{k}_1 = \mathbf{f} \left( t_n, \mathbf{y}_n + \frac{1}{4}h\mathbf{k}_1 - \frac{1}{4}h\mathbf{k}_2 \right),$$

$$\mathbf{k}_2 = \mathbf{f} \left( t_n + \frac{2}{3}h, \mathbf{y}_n + \frac{1}{4}h\mathbf{k}_1 + \frac{5}{12}h\mathbf{k}_2 \right),$$

(„zwei“ i.A. nichtlineare Gleichungen für  $\mathbf{k}_1$  und  $\mathbf{k}_2$ )

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \frac{1}{4}h(\mathbf{k}_1 + 3\mathbf{k}_2).$$

(Beispiel 2 aus Abschnitt 2.2 ist ein weiteres implizites zweistufiges RKV.)

0	0	0	0
1/2	1/2	0	0
1	-1	2	0
<hr/>			
	1/6	4/6	1/6

symbolisiert ein dreistufiges explizites RKV.

### Verfahren dritter Ordnung von Kutta:

$$k_1 = f(t_n, y_n),$$

$$k_2 = f\left(t_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_1\right),$$

$$k_3 = f(t_n + h, y_n - hk_1 + 2hk_2),$$

$$y_{n+1} = y_n + \frac{1}{6}h(k_1 + 4k_2 + k_3).$$

0	0	0	0
1/3	1/3	0	0
2/3	0	2/3	0
	1/4	0	3/4

symbolisiert ein dreistufiges explizites RKV.

### Verfahren dritter Ordnung von Heun:

$$\begin{aligned}
 k_1 &= f(t_n, y_n), \\
 k_2 &= f\left(t_n + \frac{1}{3}h, y_n + \frac{1}{3}h k_1\right), \\
 k_3 &= f\left(t_n + \frac{2}{3}h, y_n + \frac{2}{3}h k_2\right), \\
 y_{n+1} &= y_n + \frac{1}{4}h(k_1 + 3k_3).
 \end{aligned}$$

(Vgl. Beispiel 1 aus Abschnitt 2.2.)

0	0	0	0	0
1/2	1/2	0	0	0
1/2	0	1/2	0	0
1	0	0	1	0
<hr/>				
	1/6	2/6	2/6	1/6

symbolisiert ein vierstufiges explizites RKV.

### Klassisches Runge-Kutta-Verfahren:

$$k_1 = f(t_n, y_n),$$

$$k_2 = f(t_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_1),$$

$$k_3 = f(t_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_2),$$

$$k_4 = f(t_n + h, y_n + hk_3),$$

$$y_{n+1} = y_n + \frac{1}{6}h(k_1 + 2k_2 + 2k_3 + k_4).$$



Eine alternative Form von (RKV) ist

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h \sum_{j=1}^m \beta_j \mathbf{f}(t_n + \gamma_j h, \tilde{\mathbf{y}}_j) \quad (\text{RKV}^*)$$

mit  $\tilde{\mathbf{y}}_j = \mathbf{y}_n + h \sum_{\ell=1}^m \alpha_{j,\ell} \mathbf{f}(t_n + \gamma_\ell h, \tilde{\mathbf{y}}_\ell) \quad (j = 1, \dots, m).$

Setze  $\mathbf{k}_j = \mathbf{f}(t_n + \gamma_j h, \tilde{\mathbf{y}}_j)$ . Die  $\tilde{\mathbf{y}}_j$  können als Approximationen an die Lösung  $\mathbf{y}$  zur Zeit  $t_n + \gamma_j h$  interpretiert werden, die  $\mathbf{k}_j$  als Approximationen an  $\mathbf{y}'(t_n + \gamma_j h)$ .

## 4.2 Konsistenzordnung

Jedes RKV hat die Form

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h\Phi_f(\mathbf{y}_n, t_n; h) \quad \text{mit} \quad \Phi_f(\mathbf{y}_n, t_n; h) = \sum_{j=1}^m \beta_j \mathbf{k}_j.$$

Es ist ein Einschrittverfahren ( $\rho(\zeta) = \zeta - 1$ ), also stabil und (vgl. Abschnitt 2.3) genau dann konsistent, wenn  $\Phi_f(\mathbf{y}(t_n), t_n; 0) = \mathbf{f}(t_n, \mathbf{y}(t_n))\rho'(1)$  erfüllt ist, was hier zu  $\sum_{j=1}^m \beta_j = 1$  äquivalent ist.

Ein RKV ist deshalb genau dann konvergent, wenn

$$\sum_{j=1}^m \beta_j = 1$$

gilt.

Um die Konsistenzordnung eines RKVs zu bestimmen (oder um  $m$ -stufige RKV mit möglichst hoher Konsistenzordnung zu konstruieren), sind wie im Fall der Taylor-Verfahren (siehe Abschnitt 2.5) komplizierte Rechnungen erforderlich. Wir untersuchen als Beispiel explizite dreistufige RKV,

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \gamma_2 & \gamma_2 & 0 & 0 \\ \gamma_3 & \gamma_3 - \alpha_{3,2} & \alpha_{3,2} & 0 \\ \hline & \beta_1 & \beta_2 & \beta_3 \end{array},$$

und entwickeln

$$\frac{1}{h} \mathbf{R}_{n+1} = \frac{\mathbf{y}(t_{n+1}) - \mathbf{y}(t_n)}{h} - \Phi_f(\mathbf{y}(t_n), t_n; h) = \frac{\mathbf{y}(t_{n+1}) - \mathbf{y}(t_n)}{h} - \sum_{j=1}^3 \beta_j \mathbf{k}_j$$

nach Potenzen von  $h$  (unter der Voraussetzung, dass  $\mathbf{y}$  bzw.  $f$  genügend oft differenzierbar sind).

Für skalare AWPes ergibt sich mit den Abkürzungen

$$F := f_t + f_y f \quad \text{und} \quad G := f_{tt} + 2f_{ty}f + f_{yy}f^2$$

(alle Ableitungen von  $f$  werden an der Stelle  $(t_n, y(t_n))$  ausgewertet) die Beziehung

$$\frac{y(t_{n+1}) - y(t_n)}{h} = f + \frac{1}{2}Fh + \frac{1}{6}(G + f_y F)h^2 + \mathcal{O}(h^3).$$

Andererseits ist

$$k_1 = f(t_n, y(t_n)) = f,$$

$$k_2 = f(t_n + h\gamma_2, y(t_n) + h\gamma_2 k_1) = f + h\gamma_2 F + \frac{1}{2}h^2\gamma_2^2 G + \mathcal{O}(h^3),$$

$$\begin{aligned} k_3 &= f(t_n + h\gamma_3, y(t_n) + h(\gamma_3 - \alpha_{3,2})k_1 + h\alpha_{3,2}k_2) \\ &= f + h\gamma_3 F + h^2(\gamma_2\alpha_{3,2}Ff_y + \frac{1}{2}\gamma_3^2 G) + \mathcal{O}(h^3). \end{aligned}$$

Das bedeutet:

$$\begin{aligned} \frac{1}{h} \mathbf{R}_{n+1} = & \left[ 1 - \sum_{j=1}^3 \beta_j \right] f + \left[ \frac{1}{2} - \beta_2 \gamma_2 - \beta_3 \gamma_3 \right] Fh \\ & + \left[ \left( \frac{1}{3} - \beta_2 \gamma_2^2 - \beta_3 \gamma_3^2 \right) \frac{1}{2} G + \left( \frac{1}{6} - \beta_3 \gamma_2 \alpha_{3,2} \right) F f_y \right] h^2 + \mathcal{O}(h^3). \end{aligned}$$

### Folgerungen.

1. Das Euler-Verfahren ist das einzige einstufige explizite RKV der Ordnung 1 ( $\beta_1 = 1$ ). Es gibt kein einstufiges explizites RKV höherer Ordnung.
2. Die zweistufigen expliziten RKV der Ordnung 2 sind durch

$$\beta_1 + \beta_2 = 1 \quad \text{und} \quad \beta_2 \gamma_2 = \frac{1}{2}$$

charakterisiert. Beispiele sind das modifizierte ( $\beta_1 = 0, \beta_2 = 1, \gamma_2 = \frac{1}{2}$ ) und das verbesserte Euler-Verfahren ( $\beta_1 = \beta_2 = \frac{1}{2}, \gamma_2 = 1$ ). Kein explizites zweistufiges RKV besitzt die Ordnung 3.

3. Explizite dreistufige RKV der Ordnung 3 sind durch die vier Gleichungen

$$\begin{aligned}\beta_1 + \beta_2 + \beta_3 &= 1, & \beta_2\gamma_2^2 + \beta_3\gamma_3^2 &= \frac{1}{3}, \\ \beta_2\gamma_2 + \beta_3\gamma_3 &= \frac{1}{2}, & \beta_3\gamma_2\alpha_{3,2} &= \frac{1}{6}\end{aligned}$$

charakterisiert. (Man kann zeigen, dass keine dieser Methoden die Ordnung 4 besitzt.) Beispiele sind das Verfahren von Heun ( $\beta_1 = \frac{1}{4}$ ,  $\beta_2 = 0$ ,  $\beta_3 = \frac{3}{4}$ ,  $\gamma_2 = \frac{1}{3}$ ,  $\gamma_3 = \alpha_{3,2} = \frac{2}{3}$ ) und das Verfahren von Kutta ( $\beta_1 = \frac{1}{6}$ ,  $\beta_2 = \frac{2}{3}$ ,  $\beta_3 = \frac{1}{6}$ ,  $\gamma_2 = \frac{1}{2}$ ,  $\gamma_3 = 1$ ,  $\alpha_{3,2} = 2$ ).

4. Ähnliche (kompliziertere) Rechnungen zeigen, dass es eine zweiparametrische Familie expliziter vierstufiger RKV der Ordnung 4 gibt, von denen keines die Ordnung 5 besitzt. Ein Beispiel ist das klassische Runge-Kutta-Verfahren. Weitere Beispiele sind

0	0	0	0	0	
$1/3$	$1/3$	0	0	0	
$2/3$	$-1/3$	1	0	0	(3/8-Regel)
1	1	-1	1	0	
	$1/8$	$3/8$	$3/8$	$1/8$	

0	0	0	0	0	
$2/5$	$2/5$	0	0	0	
$3/5$	$-3/20$	$3/4$	0	0	(Formel von Kuntzmann).
1	$19/44$	$-15/44$	$40/44$	0	
	$55/360$	$125/360$	$125/360$	$55/360$	

Die oben beschriebene Methode, die Ordnung eines RKVs zu bestimmen, wird für Verfahren höherer Ordnung schnell unübersichtlich: Die Koeffizienten eines expliziten Verfahrens der Ordnung 3 müssen 4 Gleichungen erfüllen (s.o.), während bei einem Verfahren der Ordnung 8 bereits 200 nicht-lineare Gleichungen überprüft werden müssen. Die sog. Butcher-Theorie (vgl. J. C. Butcher, *The Numerical Analysis of Ordinary Differential Equations. Runge-Kutta and General Linear Methods*. John Wiley & Sons, Chichester 1987) erleichtert mit Hilfe graphentheoretischer Bäume die Buchhaltung bei den partiellen Ableitungen von  $f$  und erlaubt eine elegante Berechnung der Ordnung eines gegebenen RKVs (sie liefert aber keine Methode, ein Verfahren mit gewünschter Ordnung zu konstruieren).

Wir beschränken uns hier darauf, *notwendige Ordnungsbedingungen* abzuleiten, die sich aus den speziellen AWPen  $y' = y + t^{\ell-1}$ ,  $y(0) = 0$ , ( $\ell \in \mathbb{N}$ ) ergeben.



### Satz 4.1 (Notwendige Ordnungsbedingungen für RKV)

*Das durch die Butcher-Matrix*

$$\begin{array}{c|c} c & A \\ \hline & b^\top \end{array}$$

*definierte RKV besitze die Ordnung  $p$ . Dann gelten*

$$b^\top A^k C^{\ell-1} e = \frac{(\ell-1)!}{(\ell+k)!} = \frac{1}{\ell(\ell+1)\dots(\ell+k)}$$

*für  $\ell = 1, 2, \dots, p$  und  $k = 0, 1, \dots, p - \ell$ .*

*Dabei sind  $b := [\beta_1, \beta_2, \dots, \beta_m]^\top$ ,  $A := [\alpha_{j,\nu}]_{1 \leq j, \nu \leq m}$ ,  
 $C := \text{diag}(\gamma_1, \gamma_2, \dots, \gamma_m)$  und  $e := [1, 1, \dots, 1]^\top \in \mathbb{R}^m$ .*

Spezialfälle der notwendigen Bedingungen aus Satz 4.1 sind (für  $k = 0$ )

$$\mathbf{b}^\top C^{\ell-1} \mathbf{e} = \sum_{j=1}^m \beta_j \gamma_j^{\ell-1} = \frac{1}{\ell} \quad \text{für } \ell = 1, 2, \dots, p$$

sowie (für  $\ell = 1$  mit  $k \leftarrow k + 1$ )

$$\mathbf{b}^\top A^{k-1} \mathbf{e} = \frac{1}{k!} \quad \text{für } k = 1, 2, \dots, p.$$

**Bemerkung.** Ein explizites  $m$ -stufiges RKV besitzt höchstens die Konsistenzordnung  $m$ , denn hier ist  $A^m = O$  ( $A$  ist echte untere Dreiecksmatrix). Für die optimale Ordnung  $p(m)$  eines expliziten  $m$ -stufigen RKVs gilt sogar  $p(m) \leq m - 1$  falls  $m \geq 5$ , genauer:

$m$	1	2	3	4	5	6	7	8	9	10	11	12
$p(m)$	1	2	3	4	4	5	6	6	7	7	8	9

## 4.3 Absolute Stabilität

Wir wenden ein  $m$ -stufiges RKV auf die Testgleichung  $y' = \lambda y$  an und erhalten mit  $\hat{h} = \lambda h$

$$y_{n+1} = \left[ 1 + \hat{h} \mathbf{b}^\top (I_m - \hat{h} A)^{-1} \mathbf{e} \right] y_n =: R(\hat{h}) y_n,$$

so dass (bei festem  $h$ )  $\lim_{n \rightarrow \infty} y_n = 0$  (für alle  $y_0$ ) genau dann gilt, wenn  $|R(\hat{h})| < 1$  erfüllt ist. In Analogie zu Abschnitt 3.5 definieren wir den **Stabilitätsbereich eines RKVs** durch

$$\mathcal{R}_A := \{ \hat{h} \in \mathbb{C} : |R(\hat{h})| < 1 \}.$$

Für ein beliebiges  $m$ -stufiges RKV gilt

$$R(\hat{h}) = 1 + \hat{h} \mathbf{b}^\top (I_m - \hat{h} A)^{-1} \mathbf{e} = 1 + \sum_{j=1}^{\infty} \hat{h}^j \mathbf{b}^\top A^{j-1} \mathbf{e}.$$

Besitzt das Verfahren die Ordnung  $p$ , so folgt

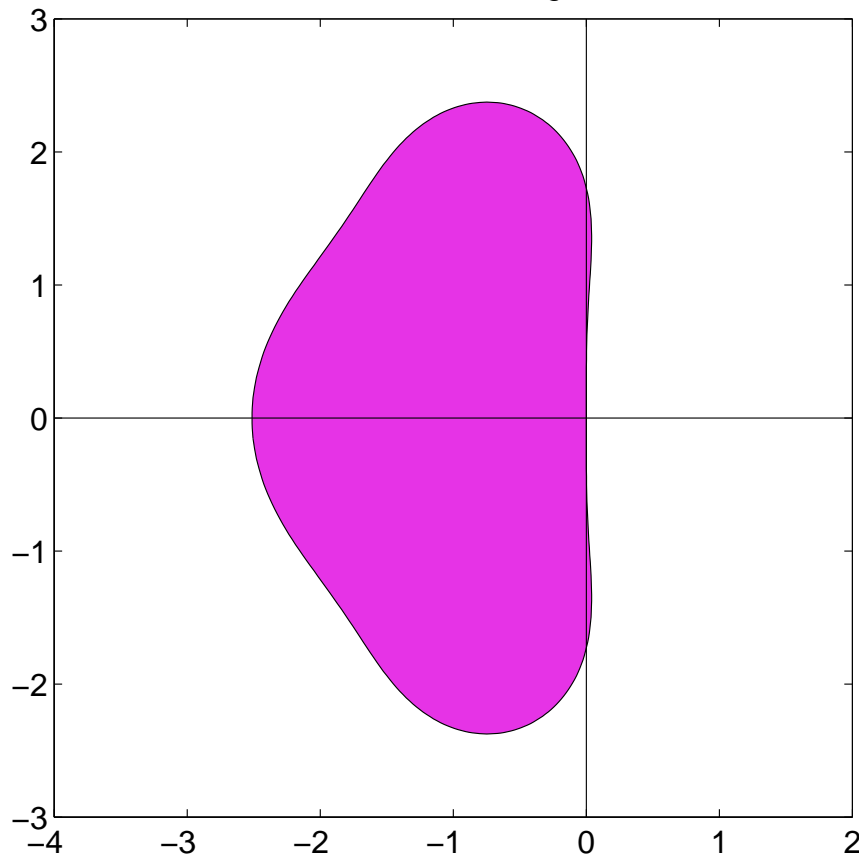
$$R(\hat{h}) = \sum_{j=0}^p \frac{1}{j!} \hat{h}^j + \sum_{j=p+1}^{\infty} \hat{h}^j \mathbf{b}^T A^{j-1} \mathbf{e}.$$

Ist das RKV explizit, so folgt

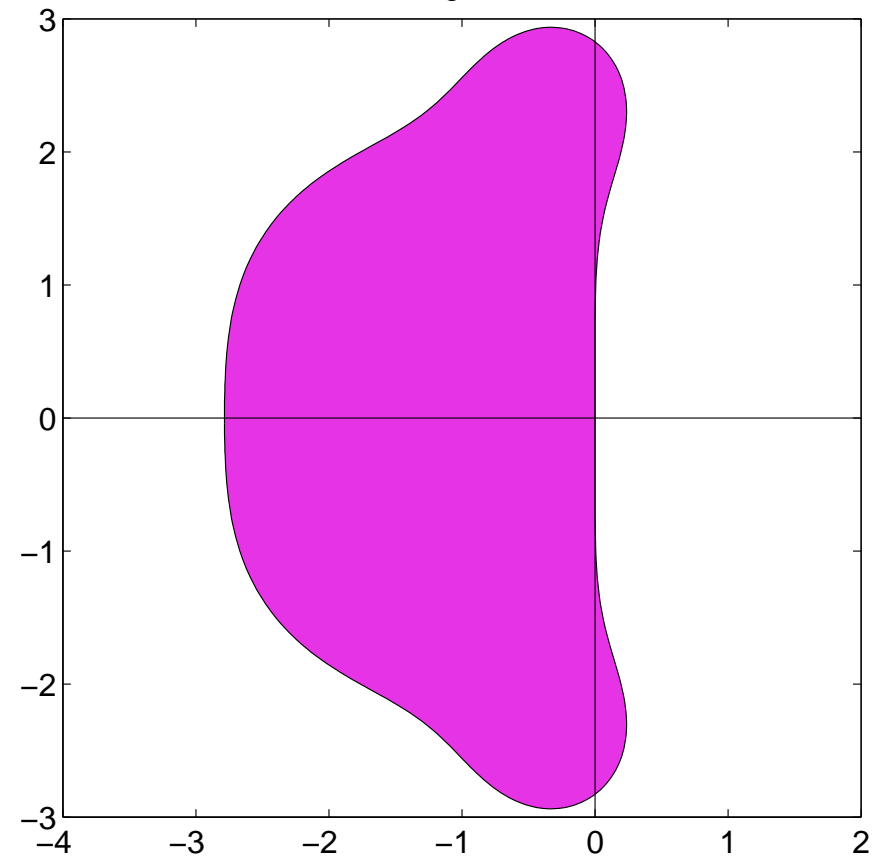
$$R(\hat{h}) = 1 + \sum_{j=1}^m \hat{h}^j \mathbf{b}^T A^{j-1} \mathbf{e}.$$

Insbesondere hängt der Stabilitätsbereich eines  $m$ -stufigen expliziten RKVs der Ordnung  $m$  ( $1 \leq m \leq 4$ ) wegen  $R(\hat{h}) = \sum_{j=0}^m \frac{1}{j!} \hat{h}^j$  nicht von den Koeffizienten des Verfahrens ab. Außerdem besitzt kein explizites RKV einen unbeschränkten Stabilitätsbereich (denn  $R$  ist ein Polynom).

Verfahren dritter Ordnung von Heun



klassisches Rung-Kutta-Verfahren



## 4.4 Eingebettete Runge-Kutta-Verfahren

Kein Verfahren zur Lösung von AWPen arbeitet in der Praxis mit einer *konstanten* Schrittweite. Man wird vielmehr versuchen, die Schrittweite an das Verhalten der Lösung  $y$  anzupassen (ändert sich  $y$  in einem Bereich schnell, so ist dort eine kleine Schrittweite angebracht; in Bereichen, in denen  $y$  kaum variiert, ist eine größere Schrittweite ausreichend). Wir werden hier eine **Schrittweitensteuerung** vorstellen, die zum Ziel hat, den Konsistenzfehler  $K_{n+1} := \frac{1}{h} R_{n+1}$  (wird in der Literatur oft lokaler Diskretisierungsfehler genannt, vgl. Abschnitt 2.3) zu kontrollieren:

$$\|K_n\| \sim \text{tol}, \quad n = 1, 2, \dots,$$

mit einer vorgegebenen Toleranz  $\text{tol}$ . Bei Systemen von DGen (insbesondere dann, wenn die Lösungskomponenten von unterschiedlicher Größenordnung sind) wird man für jede Komponente eine eigene absolute Fehlertoleranz und global eine relative Fehlertoleranz festsetzen.

Das folgende Lemma besagt, dass mit dem Konsistenzfehler auch der (eigentlich interessante) globale Diskretisierungsfehler kontrolliert wird.

### Lemma 4.2

*Für den globalen Diskretisierungsfehler  $e_n = \mathbf{y}(t_n) - \mathbf{y}_n$  eines Einschrittverfahrens gilt*

$$\|\mathbf{e}_n\| \leq (t_n - t_0) \kappa_n \exp(M(t_n - t_0)).$$

*Dabei ist  $\kappa_n := \max\{\|\mathbf{K}_j\| : j = 0, 1, \dots, n\}$  und  $M$  die Lipschitzkonstante der Verfahrensfunktion (vgl. (V<sub>2</sub>) aus Abschnitt 2.2).*

Um den Konsistenzfehler zu schätzen, verwendet man **zwei Methoden unterschiedlicher Konsistenzordnungen** (sagen wir  $p$  und  $q$  mit  $p < q$ ), um  $\mathbf{y}_n$  aus  $\mathbf{y}_{n-1}$  zu berechnen:

$$\mathbf{y}_n = \mathbf{y}_{n-1} + h\Phi_f(\mathbf{y}_{n-1}, t_{n-1}; h) \text{ bzw. } \hat{\mathbf{y}}_n = \mathbf{y}_{n-1} + h\hat{\Phi}_f(\mathbf{y}_{n-1}, t_{n-1}; h).$$

Für die zugehörigen Konsistenzfehler gelten:

$$\mathbf{K}_n = \frac{\mathbf{y}(t_n) - \mathbf{y}(t_{n-1})}{h} - \Phi(\mathbf{y}_{n-1}, t_{n-1}; h) = \mathcal{O}(h^p),$$
$$\widehat{\mathbf{K}}_n = \frac{\mathbf{y}(t_n) - \mathbf{y}(t_{n-1})}{h} - \widehat{\Phi}(\mathbf{y}_{n-1}, t_{n-1}; h) = \mathcal{O}(h^q).$$

Daraus folgt

$$\mathbf{K}_n - \widehat{\mathbf{K}}_n = \widehat{\Phi}(\mathbf{y}_{n-1}, t_{n-1}; h) - \Phi(\mathbf{y}_{n-1}, t_{n-1}; h) = \frac{1}{h}(\widehat{\mathbf{y}}_n - \mathbf{y}_n).$$

Wegen  $\mathbf{K}_n - \widehat{\mathbf{K}}_n = \mathbf{K}_n(1 + \mathcal{O}(h^{q-p})) \sim \mathbf{K}_n$  erhalten wir aus

$$\frac{1}{h} \|\mathbf{y}_n - \widehat{\mathbf{y}}_n\| \sim \|\mathbf{K}_n\|$$

eine (grobe) Schätzung für  $\|\mathbf{K}_n\|$ .



Ist  $\frac{1}{h} \|\mathbf{y}_n - \hat{\mathbf{y}}_n\| > \text{tol}$ , so wird die Schrittweite  $h$  verworfen und mit

$$\left(\frac{\tilde{h}}{h}\right)^p = \alpha \frac{h \text{ tol}}{\|\mathbf{y}_n - \hat{\mathbf{y}}_n\|} \quad (*)$$

eine neue Schrittweite  $\tilde{h}$  bestimmt ( $\alpha$  ist hier ein Sicherheitsfaktor, etwa  $\alpha = 0.9$ ). Ausgehend von  $\mathbf{y}_{n-1}$  werden jetzt neue Näherungen  $\mathbf{y}_n$  und  $\hat{\mathbf{y}}_n$  (an der Stelle  $t_{n-1} + \tilde{h}$ ) berechnet. Diesen Prozess wiederholt man so lange, bis  $\frac{1}{\tilde{h}} \|\mathbf{y}_n - \hat{\mathbf{y}}_n\| \leq \text{tol}$  erfüllt ist. Dann wird (\*) verwendet, um eine neue (größere) Schrittweite für den nächsten Schritt ( $n \rightarrow n+1$ ) vorzuschlagen.

Die Wahl von  $\tilde{h}$  nach (\*) motiviert sich folgendermaßen:

$$\begin{aligned} \text{benutzte Schrittweite } h: & \quad \frac{1}{h} \|\mathbf{y}_n - \hat{\mathbf{y}}_n\| \sim \|\mathbf{K}_n\| = ch^p + \mathcal{O}(h^{p+1}) \sim ch^p, \\ \text{erwünschte Schrittweite } \tilde{h}: & \quad \text{tol} = \|\mathbf{K}_n\| = c\tilde{h}^p + \mathcal{O}(\tilde{h}^{p+1}) \sim c\tilde{h}^p. \end{aligned}$$

Um den Aufwand in Grenzen zu halten, verwendet man zur Berechnung von  $y_n$  und  $\hat{y}_n$  zwei RKV (verschiedener Ordnungen), deren Butcher-Matrizen sich nur im Vektor  $b$  unterscheiden (d.h.  $A$  und  $c$  sind für beide Verfahren gleich, so dass die Größen  $k_j$  nur einmal berechnet werden müssen). Man spricht von **eingebetteten RKV** und schreibt

$$\begin{array}{c|c} c & A \\ \hline & b^\top \\ \hline & \hat{b}^\top \end{array}, \quad \text{z.B.} \quad \begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ \hline & 1 & 0 \\ \hline & 1/2 & 1/2 \end{array}.$$

Im Beispiel wird ein RKV der Ordnung 1 (das Euler-Verfahren) in ein RKV der Ordnung 2 (das verbesserte Euler-Verfahren) eingebettet.

Ein populäres Beispiel ist die **Fehlberg 4(5)-Formel**:

0	0	0	0	0	0	0
$\frac{1}{4}$	$\frac{1}{4}$	0	0	0	0	0
$\frac{3}{8}$	$\frac{3}{32}$	$\frac{9}{32}$	0	0	0	0
$\frac{12}{13}$	$\frac{1932}{2197}$	$-\frac{7200}{2197}$	$\frac{7296}{2197}$	0	1	0
1	$\frac{439}{216}$	-8	$\frac{3680}{513}$	$-\frac{845}{4104}$	0	0
$\frac{1}{2}$	$-\frac{8}{27}$	2	$-\frac{3544}{2565}$	$\frac{1859}{4104}$	$-\frac{11}{40}$	0
	$\frac{25}{216}$	0	$\frac{1408}{2565}$	$\frac{2197}{4104}$	$-\frac{1}{5}$	0
	$\frac{16}{135}$	0	$\frac{6656}{12825}$	$\frac{28561}{56430}$	$-\frac{9}{50}$	$\frac{2}{55}$

Hier werden zwei sechsstufige RKV der Ordnungen 4 bzw. 5 kombiniert.

## 4.5 Implizite und halb-implizite Verfahren

Ist die Matrix  $A$  eines  $m$ -stufigen RKVs keine echte untere  $\Delta$ -Matrix (ist das RKV also **implizit**), so muss in jedem Zeitschritt ein nicht-lineares Gleichungssystem der Form

$$\begin{aligned} \mathbf{k}_1 &= \mathbf{f}(t_n + \gamma_1 h, h \sum_{\ell=1}^m \alpha_{1,\ell} \mathbf{k}_\ell) \\ \vdots &\quad \quad \quad \vdots \\ \mathbf{k}_m &= \mathbf{f}(t_n + \gamma_m h, h \sum_{\ell=1}^m \alpha_{m,\ell} \mathbf{k}_\ell) \end{aligned} \quad (\square)$$

gelöst werden. Dieses System hat also  $mn$  Unbekannte, wenn  $\mathbf{y}' = \mathbf{f}(t, \mathbf{y})$  aus  $n$  Gleichungen besteht. Mit Hilfe des Banachschen Fixpunktsatzes erkennt man, dass  $(\square)$  für genügend kleine  $h$  eindeutig lösbar ist. In der Praxis wird man dieses System aber nicht mit der Fixpunktiteration, sondern mit einem Newton- bzw. Quasi-Newton-Verfahren lösen.

Ist  $A$  eine untere (aber keine echte untere)  $\Delta$ -Matrix, so nennt man das zugehörige RKV **halb-implizit**. Das System  $(\square)$  zerfällt dann in  $m$  Systeme mit jeweils  $n$  Unbekannten.

Implizite RKV werden oft mit Hilfe von **Gauß-Quadraturformeln** konstruiert. Dies sind Formeln der Bauart

$$\int_a^b g(\tau) d\tau = \sum_{j=1}^m \beta_j g(\gamma_j) + R_m(g).$$

Hier werden die Gewichte  $\beta_j$  und Knoten  $\gamma_j$  so gewählt, dass  $R_m(p) = 0$  für Polynome  $p$  möglichst hohen Grades  $d$  erfüllt ist. Man kann zeigen (vgl. Numerik I), dass eine optimale Wahl auf  $d = 2m$  führt (man sagt die Quadraturformel hat **Exaktheitsgrad**  $2m$ ).

Die zugehörigen RKV (auch sie werden **Gauß-Formeln** genannt) haben die Ordnung  $2m$ . Beachte, dass kein  $m$ -stufiges RKV eine höhere Ordnung besitzen kann. Warum?

Für  $m = 1$  ergibt sich die **implizite Mittelpunktsregel**  $\frac{1/2}{1/2} \mid \frac{1/2}{1}$ ,  
welche die Ordnung 2 besitzt.

Für  $m = 2$  und  $m = 3$  ergeben sich die Gauß-Formeln

$$\begin{array}{c|cc}
 \frac{3-\sqrt{3}}{6} & \frac{1}{4} & \frac{3-2\sqrt{3}}{12} \\
 \frac{3+\sqrt{3}}{6} & \frac{3+2\sqrt{3}}{12} & \frac{1}{4} \\
 \hline
 & \frac{1}{2} & \frac{1}{2}
 \end{array}
 \quad \text{bzw.} \quad
 \begin{array}{c|ccc}
 \frac{5-\sqrt{15}}{10} & \frac{5}{36} & \frac{10-3\sqrt{15}}{45} & \frac{25-6\sqrt{15}}{180} \\
 \frac{1}{2} & \frac{10+3\sqrt{15}}{72} & \frac{2}{9} & \frac{10-3\sqrt{15}}{72} \\
 \frac{5+\sqrt{15}}{10} & \frac{25+6\sqrt{15}}{180} & \frac{10+3\sqrt{15}}{45} & \frac{5}{36} \\
 \hline
 & \frac{5}{18} & \frac{4}{9} & \frac{5}{18}
 \end{array}$$

mit den Konsistenzordnungen 4 bzw. 6.

Bei **Gauß-Radau-Integrationsformeln** wählt man einen Knoten als (entweder linken oder rechten) Endpunkt des Integrationsintervalls.

Die übrigen Knoten und alle Gewichte werden so bestimmt, dass sich ein möglichst hoher Exaktheitsgrad ergibt. Man kann zeigen, dass eine Gauß-Radau-Formel mit  $m$  Knoten den Exaktheitsgrad  $2m - 1$  besitzt. Daher haben die zugehörigen impliziten RKV die Konsistenzordnung  $2m - 1$ . Zu dieser Klasse gehören die Verfahren

$$\begin{array}{c|c} 1 & 1 \\ \hline & 1 \end{array} \quad \text{und} \quad \begin{array}{c|cc} \frac{1}{3} & \frac{5}{12} & -\frac{1}{12} \\ 1 & \frac{3}{4} & \frac{1}{4} \\ \hline & \frac{3}{4} & \frac{1}{4} \end{array} .$$

Schließlich kann man noch beide Enden des Integrationsintervalls als Knoten wählen und die übrigen Daten so bestimmen, dass die zugehörige Integrationsformel den Exaktheitsgrad  $2m - 2$  (bzw. das zugehörige implizite RKV die Konsistenzordnung  $2m - 2$ ) besitzt. Man spricht von **Gauß-Lobatto-Formeln**. Ein Beispiel ist die Trapezregel (für  $m = 2$ ).

Ein Vorteil von impliziten gegenüber expliziten RKV ist ihr wesentlich größerer Stabilitätsbereich (wird ausführlicher im nächsten Kapitel diskutiert). Wir betrachten die Trapezregel

$$\begin{array}{c|cc}
 0 & 0 & \\
 1 & 1/2 & 1/2 \\
 \hline
 & 1/2 & 1/2
 \end{array}
 \quad \text{d.h.} \quad
 \begin{aligned}
 \mathbf{k}_1 &= \mathbf{f}(t_n, \mathbf{y}_n), \\
 \mathbf{k}_2 &= \mathbf{f}(t_{n+1}, \mathbf{y}_n + h(\mathbf{k}_1 + \mathbf{k}_2)/2), \\
 \mathbf{y}_{n+1} &= \mathbf{y}_n + h(\mathbf{k}_1 + \mathbf{k}_2)/2,
 \end{aligned}$$

oder kürzer:  $\mathbf{y}_{n+1} = \mathbf{y}_n + h(\mathbf{f}(t_n, \mathbf{y}_n) + \mathbf{f}(t_{n+1}, \mathbf{y}_{n+1}))/2$ .

Die zugehörige Stabilitätsfunktion ist  $R(\hat{h}) = (1 + \hat{h}/2)/(1 - \hat{h}/2)$  und es gilt:  $|R(\hat{h})| < 1 \Leftrightarrow |1 + \hat{h}/2| < |1 - \hat{h}/2| \Leftrightarrow \text{Real}(\hat{h}) < 0$ . Die Trapezregel ist daher **A-stabil**. (Man nennt ein Verfahren A-stabil, wenn sein Stabilitätsbereich die (offene) linke Halbebene enthält.)



## 4.6 Kollokationsmethoden

**Kollokationsmethoden** sind spezielle implizite RKV, die — auf Grund ihrer Konstruktion — sehr viel leichter zu analysieren sind als allgemeine RKV.

Mit gegebenen  $0 \leq \gamma_1 < \gamma_2 < \dots < \gamma_m \leq 1$  setzen wir  $t^{(j)} := t_n + \gamma_j h$  ( $j = 1, 2, \dots, m$ ) und suchen ein „Polynom“  $\mathbf{p}$  vom Grad  $\leq m$  (bei Systemen von  $k$  DGen ist das ein „Vektor“  $\mathbf{p} = [p_1, p_2, \dots, p_k]^\top$  aus  $k$  Polynomen vom Grad höchstens  $m$ ), das die Interpolationsbedingungen

$$\mathbf{p}(t_n) = \mathbf{y}_n, \quad \mathbf{p}'(t^{(j)}) = \mathbf{f}(t^{(j)}, \mathbf{p}(t^{(j)})) \quad (j = 1, 2, \dots, m)$$

erfüllt. Die Näherung  $\mathbf{y}_{n+1}$  an der Stelle  $t_{n+1}$  wird dann definiert durch

$$\mathbf{y}_{n+1} := \mathbf{p}(t_{n+1}).$$

$\mathbf{p}'$  ist ein Polynom vom Grad  $m - 1$ , das durch die letzten  $m$  dieser Interpolationsbedingungen eindeutig bestimmt ist. In Lagrange-Form besitzt es die Darstellung

$$\mathbf{p}'(t_n + \tau h) = \sum_{j=1}^m \ell_j(t_n + \tau h) \mathbf{k}_j$$

mit  $\ell_j(t_n + \tau h) = \prod_{j \neq i=1}^m \frac{\tau - \gamma_i}{\gamma_j - \gamma_i}$  und  $\mathbf{k}_j := \mathbf{p}'(t^{(j)})$ .

Jetzt folgt für jedes  $i \in \{1, 2, \dots, m\}$

$$\begin{aligned} \mathbf{p}(t^{(i)}) - \mathbf{p}(t_n) &= \int_{t_n}^{t^{(i)}} \mathbf{p}'(t_n + sh) ds = \int_{t_n}^{t^{(i)}} \sum_{j=1}^m \ell_j(t_n + sh) \mathbf{k}_j ds \\ &= h \sum_{j=1}^m \left( \int_0^{\gamma_i} \ell_j(r) dr \right) \mathbf{k}_j =: h \sum_{j=1}^m \alpha_{i,j} \mathbf{k}_j. \end{aligned}$$

Analog:

$$\mathbf{p}(t_{n+1}) - \mathbf{p}(t_n) = h \sum_{j=1}^m \left( \int_0^1 \ell_j(r) dr \right) \mathbf{k}_j =: h \sum_{j=1}^m \beta_j \mathbf{k}_j.$$

Daraus folgt

$$\mathbf{y}_{n+1} = \mathbf{p}(t_{n+1}) = \mathbf{p}(t_n) + h \sum_{j=1}^m \beta_j \mathbf{k}_j = \mathbf{y}_n + h \sum_{j=1}^m \beta_j \mathbf{k}_j$$

mit  $\mathbf{k}_j = \mathbf{p}'(t^{(j)}) = \mathbf{f}(t^{(j)}, \mathbf{p}(t^{(j)})) = \mathbf{f}(t_n + \gamma_j h, \mathbf{y}_n + \sum_{\ell=1}^m \alpha_{j,\ell} \mathbf{k}_\ell)$ .

Mit anderen Worten: Jedes Kollokationsverfahren ist ein (implizites) RKV. Implementiert wird es in der Form (RKV) bzw. (RKV\*), d.h. zu gegebenen  $\gamma_j$  ( $j = 1, 2, \dots, m$ ) bestimmt man zunächst

$$\alpha_{i,j} = \int_0^{\gamma_i} \ell_j(r) dr \quad \text{und} \quad \beta_j = \int_0^1 \ell_j(r) dr$$

( $i, j = 1, 2, \dots, m$ ).

Nicht jedes implizite RKV ist ein Kollokationsverfahren.

**Beispiel.**

0	0	0
2/3	1/3	1/3
	1/4	3/4

repräsentiert kein Kollokationsverfahren.

### Satz 4.3 (Konsistenzordnung bei Kollokationsverfahren)

*Für ein  $m$ -stufiges Kollokationsverfahren mit der Butcher-Matrix*

$$\begin{array}{c|c} c & A \\ \hline & b^\top \end{array}$$

*sind die folgenden drei Aussagen äquivalent:*

- *Das Verfahren besitzt die Konsistenzordnung  $m + p$ .*
- *$\int_0^1 \tau^j \prod_{j=1}^m (\tau - \gamma_j) d\tau = 0$  für  $j = 0, 1, \dots, p - 1$ .*
- *$b^\top C^{\ell-1} e = 1/\ell$  für  $\ell = 1, 2, \dots, m + p$ .*

*Dabei sind  $C := \text{diag}(\gamma_1, \gamma_2, \dots, \gamma_m)$  und  $e := [1, 1, \dots, 1]^\top \in \mathbb{R}^m$ .*