

## 4 Lineare Mehrschrittmethoden

Die bisher betrachteten Differenzenformeln waren alle *Einschrittformeln*, d. h.: Bei ihnen wird der Wert  $y_n$  jeweils allein aus dem vorausgehenden  $y_{n-1}$  berechnet. Formeln, bei denen dazu auf die  $R$  vorausgehenden Werte  $y_{n-1}, \dots, y_{n-R}$  zurückgegriffen wird, heißen „Mehrschrittformeln“ bzw. „R-Schrittformeln“. Für lineare Mehrschrittmethoden schreiben wir im folgenden kurz LMM. Zur Durchführung einer solchen Methode benötigt man Startwerte  $y_0, \dots, y_{R-1}$ , die etwa durch eine vorgeschaltete Einschrittmethode ausreichend hoher Ordnung berechnet werden. Der Mehrschrittansatz bietet im im Gegensatz zu den Einschrittformeln den Vorteil der leichten Konstruktion von Formeln beliebig hoher Ordnung und einer recht einfachen Verfahrensstruktur. Zur Vereinfachung der Notation beschränken wir uns im folgenden auf die Betrachtung von äquidistanten Zeitgittern  $\{t_n = t_0 + nh, n \geq 0\}$ . In der Praxis müssen natürlich auch variable Schrittweiten zugelassen werden, was bei Mehrschrittformeln aber gewisse technische Schwierigkeiten mit sich bringt (s. unten die diesbezügliche Bemerkung im Zusammenhang mit der Schrittweitensteuerung).

### 4.1 Konstruktion linearer Mehrschrittformeln

Mehrschrittformeln lassen sich z. B. durch numerische Integration erzeugen. Die Lösung  $u$  der AWA genügt der Integralbeziehung

$$u(t_n) = u(t_{n-\sigma}) + \int_{t_{n-\sigma}}^{t_n} f(s, u(s)) ds,$$

für festes  $\sigma \in \mathbb{N}$ . Das Integral auf der rechten Seite wird durch eine Quadraturformel approximiert. Dazu wird die Funktion  $f(t, u(t))$  durch ein Polynom  $p_m(t)$  vom Grade  $m$  in den Gitterpunkten  $t_{k-\mu}$ ,  $0 \leq \mu \leq m$ , interpoliert:  $p_m(t_{k-\mu}) = f(t_{k-\mu}, u(t_{k-\mu}))$ ,  $0 \leq \mu \leq m$ . In der Lagrange-Darstellung ist dieses Polynom gegeben durch

$$p_m(t) = \sum_{\mu=0}^m f(t_{k-\mu}, u(t_{k-\mu})) L_{\mu}^{(m)}(t), \quad L_{\mu}^{(m)}(t) = \prod_{l=0, l \neq \mu}^m \frac{t - t_{k-l}}{t_{k-\mu} - t_{k-l}},$$

und der Interpolationsfehler hat die Darstellung

$$f(t, u(t)) - p_m(t) = \frac{L(t)}{(m+1)!} f^{(m+1)}(\xi_t, u(\xi_t)) = \frac{L(t)}{(m+1)!} u^{(m+2)}(\xi_t),$$

mit einer Zwischenstelle  $\xi_t \in [t_{k-m}, t_k]$  und

$$L(t) = \prod_{l=0}^m (t - t_{k-l}) \xi_t.$$

Damit erhalten wir die Beziehung

$$u(t_n) = u(t_{n-\sigma}) + \sum_{\mu=0}^m f(t_{k-\mu}, u(t_{k-\mu})) \int_{t_{n-\sigma}}^{t_n} L_{\mu}^{(m)}(s) ds + O(h^{m+2}).$$

Durch Wahl von  $\sigma \in \mathbb{N}$  und  $k \in \{n-\sigma, \dots, n\}$  ergeben sich daraus auf einem äquidistanten Gitter die folgenden *linearen* Mehrschrittformeln:

$$y_n = y_{n-\sigma} + h \sum_{\mu=0}^m \beta_{\mu} f_{k-\mu}, \quad n \geq k,$$

mit den Abkürzungen  $f_n = f(t_n, y_n)$  und

$$\beta_{\mu} = h^{-1} \int_{t_{n-\sigma}}^{t_n} L_{\mu}^{(m)}(t) dt$$

**Beispiele:** Durch numerische Integration gewonnene LMM

1. *Adams<sup>1</sup>-Bashforth<sup>2</sup>-Formeln:*  $\sigma = 1, k = n - 1$  (explizit)

$$y_n = y_{n-1} + \sum_{\mu=0}^m \underbrace{f(t_{n-1-\mu}, y_{n-1-\mu})}_{f_{n-1-\mu}} \int_{t_{n-1}}^{t_n} L_{\mu}^{(m)}(t) dt, \quad n \geq m + 1.$$

$$m = 0: \quad y_n = y_{n-1} + h f_{n-1} \quad \text{Polygonzugmethode}$$

$$m = 1: \quad y_n = y_{n-1} + \frac{1}{2} h \{3f_{n-1} - f_{n-2}\}$$

$$m = 2: \quad y_n = y_{n-1} + \frac{1}{12} h \{23f_{n-1} - 16f_{n-2} + 5f_{n-3}\}$$

$$m = 3: \quad y_n = y_{n-1} + \frac{1}{24} h \{55f_{n-1} - 59f_{n-2} + 37f_{n-3} - 9f_{n-4}\}.$$

2. *Adams-Moulton<sup>3</sup>-Formeln:*  $\sigma = 1, k = n$  (implizit)

$$y_n = y_{n-1} + \sum_{\mu=0}^m f_{n-\mu} \int_{t_{n-1}}^{t_n} L_{\mu}^{(m)}(t) dt, \quad n \geq m.$$

---

<sup>1</sup>John Couch Adams (1819–1892): Englischer Mathematiker und Astronom; Professor in Cambridge; bekannt u. a. für die theoretische Vorhersage der Existenz des Planeten Neptun auf Basis beobachteter Störungen der Umlaufbahn von Uranus.

<sup>2</sup>Francis Bashforth (1819–1912): Englischer angewandter Mathematiker; zeitweise Professor an der Royal Military Academy, Woolwich; arbeitete u. a. über Fragen der Ballistik und Oberflächenspannung; er verwendete die nach ihm benannten Adams-Bashforth-Methoden zusammen mit Adams 1883 bei der Untersuchung der Tropfenbildung.

<sup>3</sup>Forest Ray Moulton (1872–1952): US-Amerikanischer Astronom; wirkte an der University of Chicago; Modelle zur Planetenentstehung; schlug die Verwendung von Paaren expliziter und impliziter LMM im Rahmen eines Prädiktor-Korrektor-Verfahrens vor; nach ihm benannt sind neben den „AdamsMoulton-Methoden“ auch noch der „Moulton-Mondkrater“ und die „Moulton-Ebene“ in Geometrie.

$$\begin{aligned}
m = 0: & \quad y_n = y_{n-1} + hf_n \quad (\text{implizite Euler-Methode}) \\
m = 1: & \quad y_n = y_{n-1} + \frac{1}{2}h \{f_n + f_{n-1}\} \quad (\text{Trapezregel}) \\
m = 2: & \quad y_n = y_{n-1} + \frac{1}{12}h \{5f_n + 8f_{n-1} - f_{n-2}\} \\
m = 3: & \quad y_n = y_{n-1} + \frac{1}{24}h \{9f_n + 19f_{n-1} - 5f_{n-2} + f_{n-3}\}.
\end{aligned}$$

3. *Nyström*<sup>4</sup>-Formeln:  $\sigma = 2, k = n - 1$  (explizit)

$$y_n = y_{n-2} + \sum_{\mu=0}^m f_{n-1-\mu} \int_{t_{n-2}}^{t_n} L_{\mu}^{(m)}(t) dt, \quad n \geq m.$$

$$m = 0: \quad y_n = y_{n-2} + 2hf_{n-1} \quad (\text{Mittelpunktsregel})$$

4. *Milne*<sup>5</sup>-*Simpson*<sup>6</sup>-Formeln:  $\sigma = 2, k = n$  (implizit)

$$y_n = y_{n-2} + \sum_{\mu=0}^m f_{n-\mu} \int_{t_{n-2}}^{t_n} L_{\mu}^{(m)}(t) dt, \quad n \geq m.$$

$$m = 2: \quad y_n = y_{n-2} + \frac{1}{3}h \{f_n + 4f_{n-1} + f_{n-2}\} \quad (\text{Simpson-Regel}).$$

Weitere Mehrschrittformeln lassen sich mit Hilfe der numerischen Differentiation gewinnen. In der Differentialgleichung  $u'(t) = f(t, u(t))$  wird  $u(t)$  durch das Interpolationspolynom  $p_m(t)$   $m$ -ter Ordnung zu den Funktionswerten  $p_m(t_{k-\mu}) = u(t_{k-\mu}), 0 \leq \mu \leq m$ , ersetzt. Mit Hilfe der Darstellung des Interpolationsfehlers, angewendet für  $u(t)$ ,

$$u(t) - p_m(t) = \frac{L(t)}{(m+1)!} u^{(m+1)}(\xi_t), \quad \xi_t \in [t_{k-m}, t_k],$$

ergibt sich

$$\sum_{\mu=0}^m L_{\mu}^{(m)'}(t_n) u(t_{k-\mu}) = f(t_n, u(t_n)) + O(h^{m+1}).$$

Für  $t_k = t_n$  erhält man z. B. die impliziten „Rückwärtsdifferenzenformeln“ (BDF-Methoden)

$$\sum_{\mu=0}^m L_{\mu}^{(m)'}(t_n) y_{n-\mu} = f_n, \quad n \geq m.$$

<sup>4</sup>Evert Johannes Nyström (1895–1960): Finnischer angewandter Mathematiker; Schüler von Lindelöf; Professor an der TU Helsinki; numerische Methoden für Integralgleichungen.

<sup>5</sup>William Edmund Milne (1905–1981): US-Amerikanischer Mathematiker; Professor an der Oregon State University; Pionier der Numerischen Mathematik.

<sup>6</sup>Thomas Simpson (1710–1761): Englischer Mathematiker; lehrte Mathematik an der Royal Military Academy, Woolwich, und war Mitglied der Royal Society of London; veröffentlichte eine Reihe von Lehrbüchern; die nach ihm benannte „Milne“-Regel war bereits 100 Jahre vor ihm bekannt, wurde aber wohl durch seine Bücher verbreitet.

**Beispiele:** Durch numerische Differentiation gewonnene LMM (sog. „Rückwärtsdifferenzen-Formeln“ („BDF“ = „backward differencing formulas“))

$$\begin{aligned} m = 1 : \quad & y_n - y_{n-1} = hf_n \quad \text{implizite Euler-Formel} \\ m = 2 : \quad & y_n - \frac{4}{3}y_{n-1} + \frac{1}{3}y_{n-2} = \frac{2}{3}hf_n \\ m = 3 : \quad & y_n - \frac{18}{11}y_{n-1} + \frac{9}{11}y_{n-2} - \frac{2}{11}y_{n-3} = \frac{6}{11}hf_n \\ m = 4 : \quad & y_n - \frac{48}{25}y_{n-1} + \frac{36}{11}y_{n-2} - \frac{16}{11}y_{n-3} + \frac{3}{11}y_{n-4} = \frac{12}{11}hf_n \end{aligned}$$

Durch Kombination der bisher angegebenen Differenzenformeln lassen sich fast beliebig viele weitere konstruieren. Die allgemeine Form einer linearen  $R$ -Schritt-Formel ist

$$\sum_{r=0}^R \alpha_{R-r} y_{n-r} = h \sum_{r=0}^R \beta_{R-r} f_{n-r}, \quad f_m = f(t_m, y_m), \quad (4.1.1)$$

mit Konstanten  $\alpha_R = 1$  (Normalisierungskonvention; alternativ  $\sum_{r=0}^R \beta_r = 1$ ) und  $|\alpha_0| + |\beta_0| \neq 0$ . Im Falle  $\beta_R = 0$  ist die Formel *explizit*, sonst *implizit*.

**Bemerkung 4.1:** i) Die oben angegebenen Mehrschrittformeln lassen sich ohne Schwierigkeiten auch auf Systeme von gewöhnlichen Differentialgleichungen anwenden. Dabei sind lediglich  $y_n$  und  $f_n = f(t_n, y_n)$  als Vektoren zu verstehen; die Koeffizienten bleiben dieselben. Zur Anwendung einer impliziten Formel ist in diesem Fall in jedem Zeitschritt ein  $d$ -dimensionales System nichtlinearer algebraischer Gleichungen zu lösen.

ii) Die Verwendung äquidistant verteilter Interpolationspunkte bei der Konstruktion der LMM ist nicht zwingend. Man kann diese auch für nicht äquidistante Verteilungen erzeugen, was dann natürlich zu schrittweiten-abhängigen Koeffizienten führt, wodurch Notation und Konvergenzanalyse erschwert werden. Solche „nicht-äquidistante“ LMM finden wegen ihrer größeren Flexibilität, insbesondere in Verbindung mit adaptiver Zeitschrittwahl, in der Praxis Verwendung. Die allgemeine LMM zu variablen Schrittweiten hat dann die Gestalt

$$\sum_{r=0}^R \alpha_{R-r}^n y_{n-r} = h_n \sum_{r=0}^R \beta_{R-r}^n f_{n-r}, \quad f_m = f(t_m, y_m), \quad (4.1.2)$$

mit Koeffizienten, die vom jeweiligen Auswertungspunkt  $t_n$  und den verwendeten variierenden Schrittweiten  $h_{n-r}$  abhängen.

## 4.2 Stabilität und Konvergenz

Den *lokalen Diskretisierungsfehler* (Abschneidefehler) der LMM erhält man analog wie bei den Einschrittverfahren durch Einsetzen der exakten Lösung in die Differenzen-  
gleichung:

$$\tau_n^h = \tau^h(t_n) := h^{-1} \sum_{r=0}^R \alpha_{R-r} u_{n-r} - \sum_{r=0}^R \beta_{R-r} f(t_{n-r}, u_{n-r}),$$

mit der abgekürzten Schreibweise  $u_n = u(t_n)$ . Der Abschneidefehler wird auch „lokaler Diskretisierungsfehler“ genannt. Dies ist gerechtfertigt aufgrund der folgenden Aussage:

**Hilfssatz 4.1 (Lokaler Diskretisierungsfehler):** *Unter der Annahme exakter Startwerte  $y_{n-r} = u_{n-r}$  ( $r=1, \dots, R$ ) gilt für eine allgemeine LMM die Beziehung*

$$u_n - y_n = \{I - \mathcal{O}(h|\beta_R|L)\}h\tau_n^h. \quad (4.2.3)$$

Für eine explizite LMM, d. h. für  $\beta_R = 0$ , gilt demnach sogar:

$$u_n - y_n = h\tau_n^h. \quad (4.2.4)$$

**Beweis:** Für den Fehler  $e_n := u_n - y_n$  gilt aufgrund der Annahme  $e_{n-r} = 0$  ( $r=1, \dots, R$ ) und bei Beachtung von  $\alpha_R = 1$ :

$$h\tau_n^h = e_n - h\beta_R\{f(t_n, u_n) - f(t_n, y_n)\}.$$

Mit der Lipschitz-Konstante  $L$  von  $f(t, \cdot)$  folgt daher

$$\|h\tau_n^h - e_n\| \leq h|\beta_R|L\|e_n\|,$$

was die behauptete Beziehung impliziert.

Q.E.D.

**Definition 4.1 (Konsistenz):** *Die LMM heißt „konsistent“ (mit der AWA), wenn*

$$\max_{t_n \in I} \|\tau_n^h\| \rightarrow 0 \quad (h \rightarrow 0),$$

und „von der Ordnung  $p > 0$ “, wenn für hinreichend glatte Lösung  $u$  gilt

$$\max_{t_n \in I} \|\tau_n^h\| = O(h^p).$$

**Hilfssatz 4.2 (Abschneidefehler):** *Der lokale Diskretisierungsfehler einer LMM besitzt für eine analytische Lösung  $u(t)$  die Darstellung*

$$\tau^h(t) = h^{-1} \sum_{i=0}^{\infty} C_i h^i u^{(i)}(t) \quad (4.2.5)$$

mit den Koeffizienten  $C_0 = \sum_{r=0}^R \alpha_{R-r}$  und  $(0! := 1, 0^0 := 1)$

$$C_i = (-1)^i \left\{ \frac{1}{i!} \sum_{r=0}^R r^i \alpha_{R-r} + \frac{1}{(i-1)!} \sum_{r=0}^R r^{i-1} \beta_{R-r} \right\}, \quad i \in \mathbb{N}. \quad (4.2.6)$$

**Beweis:** Entwicklung von  $u_{n-r} = u(t_n - rh)$  und  $f_{n-r} = f(t_n - rh, u(t_n - rh)) = u'(t_n - rh)$  um  $h = 0$  ergibt

$$u(t_n - rh) = \sum_{i=0}^{\infty} \frac{(-rh)^i}{i!} u^{(i)}(t_n), \quad u'(t_n - rh) = \sum_{i=0}^{\infty} \frac{(-rh)^i}{i!} u^{(i+1)}(t_n).$$

Wir setzen dies in die Definition von  $\tau_n^h = \tau^h(t_n)$  ein und ordnen nach Potenzen von  $h$ :

$$\begin{aligned} \tau_n^h &= h^{-1} \sum_{r=0}^R \alpha_{R-r} u_{n-r} - \sum_{r=0}^R \beta_{R-r} f(t_{n-r}, u_{n-r}) \\ &= h^{-1} \sum_{r=0}^R \alpha_{R-r} \sum_{i=0}^{\infty} \frac{(-rh)^i}{i!} u^{(i)}(t_n) - \sum_{r=0}^R \beta_{R-r} \sum_{i=0}^{\infty} \frac{(-rh)^i}{i!} u^{(i+1)}(t_n) \\ &= h^{-1} \sum_{i=0}^{\infty} (-1)^i \sum_{r=0}^R \alpha_{R-r} \frac{r^i h^i}{i!} u^{(i)}(t_n) - h^{-1} \sum_{i=1}^{\infty} (-1)^{i-1} \sum_{r=0}^R \beta_{R-r} \frac{r^{i-1} h^i}{(i-1)!} u^{(i)}(t_n) \\ &= h^{-1} \left\{ \sum_{r=0}^R \alpha_{R-r} \right\} u(t_n) + h^{-1} \sum_{i=1}^{\infty} (-1)^i \left\{ \frac{1}{i!} \sum_{r=0}^R r^i \alpha_{R-r} + \frac{1}{(i-1)!} \sum_{r=0}^R r^{i-1} \beta_{R-r} \right\} h^i u^{(i)}(t_n). \end{aligned}$$

Koeffizientenvergleich ergibt also für  $i = 0$  und  $i \geq 1$ :

$$C_0 = \sum_{r=0}^R \alpha_{R-r}, \quad C_i = (-1)^i \left\{ \frac{1}{i!} \sum_{r=0}^R r^i \alpha_{R-r} + \frac{1}{(i-1)!} \sum_{r=0}^R r^{i-1} \beta_{R-r} \right\}.$$

Q.E.D.

**Korollar 4.1 (Konsistenzordnung):** Eine LMM ist genau dann konsistent (mit jeder AWA), wenn gilt

$$\sum_{r=0}^R \alpha_{R-r} = 0, \quad \sum_{r=0}^R r \alpha_{R-r} + \sum_{r=0}^R \beta_{R-r} = 0, \quad (4.2.7)$$

und von der (genauen) Ordnung  $p$ , wenn gilt

$$C_0 = \dots = C_p = 0, \quad C_{p+1} \neq 0. \quad (4.2.8)$$

**Beweis:** Mit Hilfe der Darstellung (4.2.5) impliziert die Konvergenz  $\tau^h(t) \rightarrow 0$  notwendig, dass  $C_0 = C_1 = 0$ . Dies ist gleichbedeutend mit (4.2.7). Für eine LMM der Ordnung  $p$  muss außerdem noch (4.2.8) erfüllt sein. Q.E.D.

Der führende Koeffizient  $C_{p+1}$  in der Entwicklung (4.2.5) für eine LMM  $p$ -ter Ordnung wird als ihre „Fehlerkonstante“ bezeichnet. Die Beziehungen (4.2.5), (4.2.6) können zur Konstruktion von LMM vorgegebener Ordnung und Struktur benutzt werden.

**Beispiel 4.1:** Die allgemeine lineare 2-Schrittmethod hat die freien Parameter

$$\alpha_0, \alpha_1, \beta_0, \beta_1, \beta_2 \quad (\alpha_2 = 1).$$

Mit  $\alpha := \alpha_0$  ergibt sich somit:

$$\begin{aligned} C_0 &= 1 + \alpha_1 + \alpha \\ C_1 &= -(\alpha_1 + 2\alpha) - (\beta_2 + \beta_1 + \beta_0) \\ C_2 &= \frac{1}{2}(\alpha_1 + 4\alpha) + (\beta_1 + 2\beta_0) \\ C_3 &= -\frac{1}{6}(\alpha_1 + 8\alpha) - \frac{1}{2}(\beta_1 + 4\beta_0) \\ C_4 &= \frac{1}{24}(\alpha_1 + 16\alpha) + \frac{1}{6}(\beta_1 + 8\beta_0) \\ C_5 &= \frac{1}{120}(\alpha_1 + 32\alpha) - \frac{1}{24}(\beta_1 + 16\beta_0) \\ &\vdots \end{aligned}$$

Zur Konstruktion von Formeln mindestens dritter Ordnung erzwingen die Bedingungen  $C_0 = \dots = C_3 = 0$ , dass

$$\alpha_1 = -1 - \alpha, \beta_0 = -\frac{1}{12}(1 + 5\alpha), \quad \beta_1 = \frac{2}{3}(1 - \alpha), \quad \beta_2 = \frac{1}{12}(5 + \alpha).$$

Die allgemeine 2-Schrittmethod (mindestens) dritter Ordnung hat also die Form

$$y_n - (1 + \alpha)y_{n-1} + \alpha y_{n-2} = \frac{1}{12}h [(5 + \alpha)f_n + 8(1 - \alpha)f_{n-1} - (1 + 5\alpha)f_{n-2}].$$

Ferner gilt:  $C_4 = -\frac{1}{4!}(1 + \alpha)$ ,  $C_5 = -\frac{1}{3 \cdot 5!}(17 + 13\alpha)$ .

(i)  $\alpha = -1$  : Es wird  $C_4 = 0$  ( $C_5 \neq 0$ ), d. h.: Die Methode ist von der Ordnung  $p = 4$  :

$$y_n = y_{n-2} + \frac{1}{3}h [f_n + 4f_{n-1} + f_{n-2}] \quad (\text{Simpson-Formel}).$$

(ii)  $\alpha \neq -1$  : Es wird  $C_4 \neq 0$ , d. h.: Die Methode ist von der Ordnung  $p = 3$ . Für  $\alpha = 0$  erhalten wir die 2-stufige implizite Adams-Moulton-Formel:

$$y_n = y_{n-1} + \frac{1}{12}h [5f_n + 8f_{n-1} - f_{n-2}].$$

(iii) Für  $\alpha = -5$  wird die Methode explizit

$$y_n + 4y_{n-1} - 5y_{n-2} = h[4f_{n-1} + 2f_{n-2}].$$

Diese explizite Formel dritter Ordnung scheint für die praktische Realisierung besonders attraktiv zu sein, da sie mit nur 2 Funktionsauswertungen pro Zeitschritt auskommt. Dies wäre wichtig bei Verwendung von adaptiv gesteuerter variabler Schrittweite, wobei unter Umständen fehlende Startwerte zusätzlich berechnet werden müssen. Die expliziten Einschrittverfahren dritter Ordnung erfordern mindestens drei Funktionsauswertungen pro Zeitschritt. Bei ihrer Anwendung auf die (monotone) AWA

$$u'(t) = -u(t), \quad t \geq 0, \quad u(0) = 1,$$

mit der Lösung  $u(t) = e^{-t}$  findet man bei 9-stelliger Rechnung und Verwendung der exakten Startwerte  $y_0 = 1, y_1 = e^{-h}$  (bis auf Maschinengenauigkeit), für den Näherungswert  $y_N \sim e^{-1} = 0.3678\dots$  die in Tabelle 4.1 angegebenen Werte.

Tabelle 4.1: Demonstration der Nichtkonvergenz einer LMM

$h$	$y_N$	$h = 0.05$	
0.2	0.398...	$y_{10}$	$0.252 \cdot 10^0$
0.1	$-0.677 \dots 10^1$	$y_{11}$	$0.240 \cdot 10^1$
0.05	$-0.465 \dots 10^7$	$y_{12}$	$-0.884 \cdot 10^1$
0.025	$-0.290 \dots 10^{18}$	$y_{13}$	$-0.489 \cdot 10^2$
		$y_{14}$	$-0.248 \cdot 10^3$
		$y_{15}$	$0.128 \cdot 10^4$
		$y_{16}$	$-0.661 \cdot 10^4$
		$y_{17}$	$0.340 \cdot 10^5$
		$y_{18}$	$-0.175 \cdot 10^6$
		$y_{19}$	$0.904 \cdot 10^6$
		$y_{20}$	$-0.465 \cdot 10^7$

Im Gegensatz zu den Einschrittmethoden, welche die Lipschitz-Bedingung ( $L_h$ ) erfüllen, ist für LMM die Konsistenz offenbar allein noch nicht hinreichend für ihre Konvergenz.

**Definition 4.2 (Konvergenz):** Eine LMM heißt „konvergent“, wenn für jede AWA gilt

$$\max_{0 \leq n \leq N} \|y_n - u(t_n)\| \rightarrow 0 \quad (h \rightarrow 0),$$

vorausgesetzt die Startwerte  $y_0, \dots, y_{R-1}$  konvergieren,

$$\max_{0 \leq n \leq R-1} \|y_n - u_0\| \rightarrow 0 \quad (h \rightarrow 0).$$

**Definition 4.3 (Charakteristische Polynome):** Für die LMM wird definiert

(i) „erstes charakteristisches Polynom“:  $\rho(\lambda) = \sum_{r=0}^R \alpha_r \lambda^r$ ,

(ii) „zweites charakteristisches Polynom“:  $\sigma(\lambda) = \sum_{r=0}^R \beta_r \lambda^r$ .

Gemäß Hilfssatz 4.1 ist die Konsistenz einer LMM äquivalent mit den Beziehungen

$$\rho(1) = 0, \quad \rho'(1) = \sigma(1).$$



**Hilfssatz 4.3 (Wurzelbedingung):** Für eine konvergente LMM gilt für die Wurzeln  $\lambda_i \in \mathbb{C}$  des ersten charakteristischen Polynoms  $\rho(\lambda)$  notwendig:

$$|\lambda_i| \leq 1, \text{ falls } \lambda_i \text{ einfach}, \quad |\lambda_i| < 1, \text{ falls } \lambda_i \text{ mehrfache.}$$

**Beweis:** Betrachte die triviale Anfangswertaufgabe  $u'(t) = 0$ ,  $u(0) = 0$ , mit der Lösung  $u \equiv 0$ . Die LMM hat dafür die Gestalt

$$\sum_{r=0}^R \alpha_{R-r} y_{n-r} = 0. \quad (4.2.9)$$

Für die Lösung dieser linearen Differenzgleichung machen wir den Ansatz  $y_n = \lambda^n$ , so dass

$$\lambda^{n-R} \sum_{r=0}^R \alpha_r \lambda^r = \lambda^{n-R} \rho(\lambda) \stackrel{!}{=} 0.$$

Die Wurzeln von  $\rho(\lambda)$  erzeugen also durch  $y_n = \lambda_i^n$  Lösungen von (4.2.9). Im Falle einer mehrfachen Wurzel  $\lambda_i$  von  $\rho(\lambda)$  ist auch durch  $y_n = n\lambda_i^n$  eine Lösung von (4.2.9) gegeben, da

$$\sum_{r=0}^R \alpha_{R-r} (n-r) \lambda_i^{n-r} = n\lambda_i^{n-R} \rho(\lambda_i) - \lambda_i^{n+1-R} \rho'(\lambda_i) = 0.$$

Seien nun o.B.d.A.  $\lambda_1$  eine einfache und  $\lambda_2$  eine mehrfache Nullstelle von  $\rho(\lambda)$ . Dann ist durch

$$y_n = h(\lambda_1^n + n\lambda_2^n)$$

sicher eine Lösung von (4.2.9) gegeben mit der speziellen Eigenschaft

$$y_n \rightarrow 0 \quad (h \rightarrow 0), \quad n = 0, \dots, R-1.$$

Aufgrund der angenommenen Konvergenz der LMM muss dann auch gelten

$$y_n = h(\lambda_1^n + n\lambda_2^n) \rightarrow 0 \quad (h \rightarrow 0, \quad t_n = t, \text{ fest}).$$

Nun ist  $t = t_n = nh$ , d. h.:

$$\frac{1}{n} \lambda_1^n + \lambda_2^n \rightarrow 0 \quad (n \rightarrow \infty).$$

Dies impliziert  $|\lambda_1| \leq 1$  und  $|\lambda_2| < 1$ .

Q.E.D.

**Definition 4.4 (Null-Stabilität):** Eine LMM heißt „null-stabil“, wenn keine Nullstelle des ersten charakteristischen Polynoms  $\rho(\lambda)$  einen Betrag größer als eins hat, und wenn alle Nullstellen mit Betrag eins einfach sind.

Die obige explizite 2-Schrittformel ist nicht null-stabil, denn ihr erstes charakteristisches Polynom

$$\rho(\lambda) = \lambda^2 + 4\lambda - 5$$

hat die Nullstellen

$$\lambda_1 = 1, \quad \lambda_2 = -5.$$

Dies erklärt ihre offensichtliche Divergenz für  $h \rightarrow 0$ .

Wir betrachten nun das gestörte Mehrschrittverfahren

$$\begin{aligned} \tilde{y}_n &= y_n + \rho_n, \quad n = 0, \dots, R-1, \\ \sum_{r=0}^R \alpha_{R-r} \tilde{y}_{n-r} &= h \sum_{r=0}^R \beta_{R-r} f(t_{n-r}, \tilde{y}_{n-r}) + \rho_n, \quad n \geq R. \end{aligned} \quad (4.2.10)$$

Sei  $L$  wieder die globale Lipschitz-Konstante der Funktion  $f(t, x)$ . Unter der Bedingung  $h < 1/(L|\beta_R|)$  (im Falle  $\beta_R \neq 0$ ) sind die Werte  $\tilde{y}_n$ ,  $n \geq 0$ , durch (4.2.10) eindeutig bestimmt (Anwendung des Banachschen Fixpunktsatzes).

**Satz 4.1 (Stabilitätssatz):** *Die AWA genüge der globalen Lipschitz-Bedingung, und die LMM sei null-stabil. Dann gilt unter der Voraussetzung  $h < 1/(L|\beta_R|)$  (im Falle  $\beta_R \neq 0$ ) für je zwei Lösungen  $\{y_n\}$  und  $\{\tilde{y}_n\}$  von (4.1.1) bzw. (4.2.10) die Abschätzung*

$$\|\tilde{y}_n - y_n\| \leq K e^{\Gamma(t_n - t_0)} \left\{ \max_{0 \leq \nu \leq R-1} \|\rho_\nu\| + \sum_{\nu=R}^n \|\rho_\nu\| \right\}, \quad n \geq R. \quad (4.2.11)$$

Die Konstanten  $K$  und  $\Gamma$  sind bestimmt durch die Lipschitz-Konstante  $L$  und die Koeffizienten  $\alpha_r, \beta_r$ ; für  $hL|\beta_R| \rightarrow 1$  gehen  $K, \Gamma \rightarrow \infty$ .

**Beweis:** Wir führen den Beweis nur für den skalaren Fall ( $d = 1$ ). Die Differenzen  $e_n = \tilde{y}_n - y_n$  genügen den Beziehungen  $e_n = \rho_n$ , für  $n=0, \dots, R-1$ , und für  $n \geq R$ :

$$\begin{aligned} e_n &= - \sum_{r=1}^R \alpha_{R-r} e_{n-r} + h\beta_R \{f(t_n, \tilde{y}_n) - f(t_n, y_n)\} + \\ &\quad + \underbrace{h \sum_{r=1}^R \beta_{R-r} \{f(t_{n-r}, \tilde{y}_{n-r}) - f(t_{n-r}, y_{n-r})\}}_{=: b_n} + \rho_n. \end{aligned}$$

Wir definieren

$$\sigma_n := \begin{cases} \frac{f(t_n, \tilde{y}_n) - f(t_n, y_n)}{e_n}, & e_n \neq 0, \\ 0, & e_n = 0. \end{cases}$$

Damit gilt dann  $|\sigma_n| \leq L$  und

$$(1 - h\beta_R\sigma_n) e_n = - \sum_{r=1}^R \alpha_{R-r} e_{n-r} + b_n. \quad n \geq R.$$

Diese Beziehung ergibt zusammen mit den trivialen Gleichungen

$$(1 - h\beta_R\sigma_{n-r})e_{n-r} = (1 - h\beta_R\sigma_{n-r})e_{n-r}, \quad r = 1, \dots, R,$$

eine rekursive Folge von Gleichungssystemen

$$D_n E_n = C_n A E_{n-1} + B_{n-1}, \quad n \geq 1,$$

für die  $(R+1)$ -Tupel

$$E_n = (e_n, e_{n-R}, \dots, e_{n-1})^T \in \mathbb{R}^{R+1}.$$

Dabei ist  $B_n = (0, \dots, 0, b_n)^T \in \mathbb{R}^{R+1}$ , und die  $(R+1) \times (R+1)$ -Matrizen  $D_n$ ,  $C_n$  und  $A$  sind definiert wie folgt:

$$D_n = \begin{bmatrix} 0 & 1-h\beta_R\sigma_{n-R} & 0 & & \\ & \ddots & \ddots & & \\ 0 & & & 1-h\beta_R\sigma_{n-1} & \\ 1-h\beta_R\sigma_n & 0 & 0 & & \end{bmatrix}, \quad C_n = \begin{bmatrix} 1-h\beta_R\sigma_{n-R} & 0 & & & \\ & \ddots & & & \\ 0 & & 1-h\beta_R\sigma_{n-1} & & \\ 0 & & & 1 & \end{bmatrix},$$

$$A = \begin{bmatrix} 0 & 1 & & 0 & \\ & \ddots & \ddots & & \\ & & 0 & 1 & \\ 0 & -\alpha_0 & & -\alpha_{R-1} & \end{bmatrix}.$$

Aufgrund der Voraussetzung  $h < 1/L|\beta_R|$  sind die Matrizen  $D_n$  invertierbar, und es gilt folglich

$$E_n = D_n^{-1}\{C_n A E_{n-1} + B_{n-1}\}, \quad n \geq 1.$$

Die Matrix  $A$  hat das charakteristische Polynom (Entwicklung nach der letzten Zeile)

$$\chi_A(\lambda) = (-1)^{R+1} (0 + \alpha_0 + \alpha_1\lambda + \dots + \alpha_{R-1}\lambda^{R-1} + \lambda^R)\lambda = (-1)^R \rho(\lambda)\lambda.$$

Aufgrund der vorausgesetzten Null-Stabilität der LMM gibt es nun gemäß Hilfssatz 4.4 (s. unten) eine natürliche Matrizenorm  $\|\cdot\|_0$ , so dass

$$\|A\|_0 = \text{spr}(A) \leq 1.$$

Die erzeugende Vektornorm sei gleichfalls mit  $\|\cdot\|_0$  bezeichnet. Damit erhalten wir die Normabschätzung

$$\|E_n\|_0 \leq \|D_n^{-1}\|_0 \{\|C_n\|_0 \|E_{n-1}\|_0 + \|B_{n-1}\|_0\}.$$

Wegen der Äquivalenz aller Normen auf dem  $\mathbb{R}^{R+1}$  gilt mit einer Konstante  $\gamma > 0$

$$\gamma^{-1} \|z\|_0 \leq \sum_{\nu=0}^R |z_\nu| \leq \gamma \|z\|_0 \quad \forall z \in \mathbb{R}^{R+1}.$$

Damit wird nun abgeschätzt:

$$\begin{aligned} |b_n| &\leq h\beta L \sum_{r=1}^R |e_{n-r}| + |\rho_n| \\ &\leq h\beta L \gamma \|E_n\|_0 + |\rho_n| \end{aligned}$$

mit  $\beta := \max_{r=0, \dots, R-1} |\beta_r|$ , und folglich:

$$\|B_n\|_0 \leq \gamma |b_n| \leq h\beta L \gamma^2 \|E_n\|_0 + \gamma |\rho_n|.$$

Weiter gilt aufgrund der Identität

$$\frac{1}{1-a} = 1 + \frac{a}{1-a}, \quad |a| < 1,$$

die Darstellung

$$D_n^{-1} = P_n^{-1} \left( I + \sum_n \right), \quad \sum_n := \text{diag}_{r=0, \dots, R-1} \left( \frac{h\beta_R \sigma_{n-r}}{1-h\beta_R \sigma_{n-r}} \right),$$

mit einer Permutationsmatrix  $P_n$ . Es ist also ( $\|P_n\|_0 = 1$ )

$$\begin{aligned} \|D_n^{-1}\|_0 &\leq 1 + \left\| \sum_n \right\|_0 \leq 1 + \gamma^2 \max_{r=0, \dots, R-1} \left\{ \left| \frac{h\beta_R \sigma_{n-r}}{1-h\beta_R \sigma_{n-r}} \right| \right\} \\ &\leq 1 + h \frac{\gamma^2 |\beta_R| L}{1-h|\beta_R|L}. \end{aligned}$$

Auf analoge Weise erhalten wir

$$\|C_n\|_0 \leq 1 + h|\beta_R|L \gamma^2.$$

Kombination aller dieser Abschätzungen ergibt nun

$$\begin{aligned} \|E_n\|_0 &\leq \left( 1 + h \frac{\gamma^2 |\beta_R| L}{1-h|\beta_R|L} \right) \{ (1 + h|\beta_R|L \gamma^2) \|E_{n-1}\|_0 + \\ &\quad + h\beta L \gamma^2 \|E_{n-1}\|_0 + \gamma |\rho_n| \} \\ &\leq \|E_{n-1}\|_0 + h\Gamma \|E_{n-1}\|_0 + \Lambda |\rho_n| \end{aligned}$$

mit den Konstanten

$$\begin{aligned} \Gamma &= \frac{\gamma^2 |\beta_R| L}{1-h|\beta_R|L} (1 + h\gamma^2 |\beta_R| L) + \gamma^2 L (|\beta_R| + \beta), \\ \Lambda &= \gamma \left( 1 + h \frac{\gamma^2 |\beta_R| L}{1-h|\beta_R|L} \right). \end{aligned}$$

Durch rekursive Anwendung dieser Ungleichung für  $n, n-1, \dots, 0$  erhalten wir

$$\|E_n\|_0 \leq \Gamma \sum_{\nu=0}^{n-1} h \|E_\nu\|_0 + \|E_0\|_0 + \Lambda \sum_{\nu=1}^n |\rho_\nu|.$$

Hierauf lässt sich nun das diskrete Gronwallsche Lemma anwenden mit dem Resultat

$$\|E_n\|_0 \leq e^{\Gamma(t_n - t_0)} \left\{ \|E_0\|_0 + \Lambda \sum_{\nu=1}^n |\rho_\nu| \right\}.$$

Dies impliziert dann schließlich mit  $K = \gamma \max\{\gamma, \Lambda\}$ :

$$|e_n| \leq K e^{\Gamma(t_n - t_0)} \left\{ \max_{\nu=0, \dots, R-1} |\rho_\nu| + \sum_{\nu=R}^n |\rho_\nu| \right\}.$$

Q.E.D.

**Hilfssatz 4.4 (Spektralradius):** *Zu jeder Matrix  $A \in \mathbb{R}^{m \times m}$  gibt es für jedes  $\varepsilon > 0$  eine natürliche Matrixnorm  $\|\cdot\|_{A,\varepsilon}$ , so dass für den Spektralradius  $\text{spr}(A)$  gilt:*

$$\text{spr}(A) \leq \|A\|_{A,\varepsilon} \leq \text{spr}(A) + \varepsilon. \quad (4.2.12)$$

*Ist jeder Eigenwert von  $A$  mit  $|\lambda| = \text{spr}(A)$  nur einfache Nullstelle des charakteristischen Polynoms  $\chi_A(z) = \det(A - zI)$ , so existiert sogar eine natürliche Matrixnorm  $\|\cdot\|_{A,0}$ , so dass*

$$\text{spr}(A) = \|A\|_{A,0}. \quad (4.2.13)$$

**Beweis:** Die Matrix  $B$  ist ähnlich zu einer Dreiecksmatrix

$$B = T^{-1}RT, \quad R = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ 0 & & r_{nn} \end{bmatrix},$$

mit den Eigenwerten von  $B$  auf der Hauptdiagonalen, d. h.:

$$\text{spr}(B) = \max_{1 \leq i \leq n} |r_{ii}|.$$

Für ein beliebiges  $\delta \in (0, 1]$  setzen wir

$$S_\delta = \begin{bmatrix} 1 & & & & 0 \\ & \delta & & & \\ & & \ddots & & \\ & & & \delta^{n-1} & \\ 0 & & & & \delta^{n-1} \end{bmatrix}, \quad R_0 = \begin{bmatrix} r_{11} & & & & 0 \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & r_{nn} \end{bmatrix},$$

$$Q_\delta = \begin{bmatrix} 0 & r_{12} & \delta r_{13} & \cdots & \delta^{n-2} r_{1n} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \delta r_{n-2,n} \\ & & & \ddots & r_{n-1,n} \\ & & & & 0 \end{bmatrix},$$

und haben damit

$$R_\delta := S_\delta^{-1}RS_\delta = \begin{bmatrix} r_{11} & \delta r_{12} & \cdots & \delta^{n-1}r_{1n} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \delta r_{n-1,n} \\ 0 & & & r_{nn} \end{bmatrix} = R_0 + \delta Q_\delta.$$

Wegen der Regularität von  $S_\delta^{-1}T$  wird durch

$$\|x\|_\delta := \|S_\delta^{-1}Tx\|_2, \quad x \in \mathbb{R}^n,$$

eine Vektornorm erklärt. Dann ist wegen  $R = S_\delta R_\delta S_\delta^{-1}$ :

$$B = T^{-1}RT = T^{-1}S_\delta R_\delta S_\delta^{-1}T$$

für alle  $x \in \mathbb{R}^n$  und  $y = S_\delta^{-1}Tx$ :

$$\begin{aligned} \|Bx\|_\delta &= \|T^{-1}S_\delta R_\delta S_\delta^{-1}Tx\|_\delta = \|R_\delta y\|_2 \\ &\leq \|R_0 y\|_2 + \delta \|Q_\delta y\|_2 \leq \{\max_{1 \leq i \leq n} |r_{ii}| + \delta \mu\} \|y\|_2 \\ &\leq \{\text{spr}(B) + \delta \mu\} \|x\|_\delta \end{aligned}$$

mit der Konstanten

$$\mu = \left( \sum_{i,j=1}^n |r_{ij}|^2 \right)^{1/2}.$$

Also ist

$$\sup_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\|Bx\|_\delta}{\|x\|_\delta} \leq \text{spr}(B) + \mu \delta,$$

und die Behauptung folgt mit  $\delta = \varepsilon/\mu$ .

Q.E.D.

Wie bei den Einschrittverfahren führt uns der Stabilitätssatz nun auch zu einem allgemeinen Konvergenzresultat für LMM.

**Satz 4.2 (Konvergenzsatz):** *Die AWA genüge der globalen Lipschitz-Bedingung, und die LMM sei null-stabil. Unter den Bedingungen*

$$h < \frac{1}{|\beta_R|L} \quad (\text{im Falle } \beta_R \neq 0), \quad (4.2.14)$$

$$\delta_h := \max_{0 \leq n \leq R-1} \|y_n - u(t_n)\| \rightarrow 0 \quad (h \rightarrow 0) \quad (4.2.15)$$

ist dann ihre Konsistenz hinreichend für die Konvergenz

$$\max_{0 \leq n \leq N} \|y_n - u(t_n)\| \rightarrow 0 \quad (h \rightarrow 0).$$

und es gilt die a priori Fehlerabschätzung

$$\|y_n - u(t_n)\| \leq K e^{\Gamma(t_n - t_0)} \left\{ \delta_h + (t_n - t_0) \max_{R \leq \nu \leq n} \|\tau_\nu^h\| \right\} \quad (4.2.16)$$

mit den Konstanten  $K, \Gamma$  aus Satz 4.1.

**Beweis:** Die exakte Lösung  $u(t)$  genügt der „gestörten“ Differenzgleichung

$$u_n = y_n + (u_n - y_n), \quad n = 0, \dots, R-1,$$

$$\sum_{r=0}^R \alpha_{R-r} u_{n-r} = h \sum_{r=0}^R \beta_{R-r} f(t_{n-r}, u_{n-r}) + h\tau_n, \quad n \geq R,$$

wobei  $u_n := u(t_n)$  gesetzt ist. Der Stabilitätssatz 4.1 liefert also für  $h < 1/(|\beta_R|L)$ :

$$\|u_n - y_n\| \leq K e^{\Gamma(t_n - t_0)} \left\{ \max_{\nu=0, \dots, R-1} \|u_\nu - y_\nu\| + h \sum_{\nu=R}^n \|\tau_\nu\| \right\}.$$

Q.E.D.

Bei den null-stabilen LMM ist also ebenfalls die *globale* Konvergenzordnung gleich der *lokalen* Konsistenzordnung. Als Anwendung des Konvergenzsatzes haben wir den folgenden Satz:

**Satz 4.3 (Konvergenz der Adams-Verfahren):** *Die LMMn vom Adams-Bashforth-, Adams-Moulton-, Nyström- und Milne-Simpson-Typ sind konvergent.*

**Beweis:** Alle diese Formeln wurden konstruiert durch Ansatz einer numerischen Integrationsformel in der Beziehung

$$u(t_n) = u(t_{n-\sigma}) + \int_{t_{n-\sigma}}^{t_n} f(t, u(t)) dt \sim u(t_{n-\sigma}) + \int_{t_{n-\sigma}}^{t_n} p_m(t) dt,$$

mit gewissen Interpolationspolynomen  $p_m(t)$  zu  $f(t, u(t))$  vom Grad  $m$ . Diese Formeln sind exakt, wenn  $u(t)$  selbst ein Polynom vom Grade  $m+1$  ist. Folglich sind diese Methoden mindestens von der Ordnung  $m+1$ , d. h. insbesondere konsistent. Ihre ersten charakteristischen Polynome  $\rho(\lambda)$  haben die Form

$$\rho(\lambda) = \lambda^k - \lambda^{k-1} \quad \text{bzw.} \quad \rho(\lambda) = \lambda^k - \lambda^{k-2}.$$

Die Methoden sind also auch null-stabil. Dasselbe gilt auch für die *Rückwärtsdifferenzformeln* bis zur Stufe  $R=6$ ; ab  $R=7$  geht die Nullstabilität verloren. Q.E.D.

Bei der Konstruktion einer  $R$ -Schrittmethode stehen  $2R+1$  freie Parameter  $\alpha_r, \beta_r$ ,  $r = 0, \dots, R$  zur Verfügung, um möglichst viele der Koeffizienten  $C_i$  in der Entwicklung des (lokalen) Abschneidefehlers zu Null zu machen ( $\alpha_R = 1!$ ). Im expliziten Fall sind es  $2R$  Parameter. Folglich ist die höchste erzielbare Ordnung  $m = 2R$  für eine implizite und  $m = 2R - 1$  für eine explizite Formel. Für konvergente Formeln besteht jedoch die folgende Beschränkung:

**Satz 4.4 (Ordnungsbarriere):** *Keine null-stabile  $R$ -Schrittmethod kann eine Ordnung  $m > R + 2$  für  $R$  gerade und  $m > R + 1$  für  $R$  ungerade haben. Für explizite  $R$ -Schrittmethoden ist die maximale Ordnung  $m = R$ .*

**Beweis:** Siehe G. Dahlquist: Convergence and stability in the numerical integration of ordinary differential equations, Math Scand. 4, 33-53 (1956). Q.E.D.

Null-stabile  $R$ -Schrittmethoden der Ordnung  $m = R + 2$  werden als „optimal“ bezeichnet. Offenbar ist die Simpson-Formel eine optimale 2-Schrittformel

$$y_n = y_{n-2} + \frac{1}{3}h\{f_n + 4f_{n-1} + f_{n-2}\}.$$

### 4.3 Numerische Stabilität linearer Mehrschrittmethoden

Wir studieren nun die numerischen Stabilitätseigenschaften der allgemeinen LMM

$$\sum_{r=0}^R \alpha_{R-r} y_{n-r} = h \sum_{r=0}^R \beta_{R-r} f_{n-r} \quad (4.3.17)$$

mit den üblichen Konventionen  $\alpha_R = 1$ ,  $|\alpha_0| + |\beta_0| \neq 0$ . Anwendung von (4.3.17) auf die Testgleichung  $u'(t) = qu(t)$  ergibt die Differenzgleichung

$$\sum_{r=0}^R [\alpha_{R-r} - hq\beta_{R-r}] y_{n-r} = 0. \quad (4.3.18)$$

Wir haben schon bei der Analyse der Konvergenzfrage für die hier betrachteten Methoden gesehen, dass (4.3.18) gelöst wird durch die Folgen  $(\lambda_i^n)_{n \in \mathbb{N}}$  bzw.  $(n\lambda_i^n)_{n \in \mathbb{N}}$  mit den einfachen bzw. mehrfachen Wurzeln  $\lambda_i$  des sog. „Stabilitätspolynoms“

$$\pi(\lambda; qh) := \sum_{r=0}^R [\alpha_r - qh\beta_r] \lambda^r = \rho(\lambda) - qh\sigma(\lambda).$$

Offensichtlich bleibt die Lösung  $y_n$  von (4.3.18) nur dann beschränkt, wenn  $|\lambda_i| \leq 1$  im Falle einer einfachen Wurzel und  $|\lambda_i| < 1$  im Falle einer mehrfachen Wurzel von  $\pi(\lambda; qh)$ . Dies legt für LMMn die folgende Definition nahe:

**Definition 4.5 (Absolute Stabilität):** *Eine LMM heißt „absolut stabil“ für ein  $\bar{h} = qh \neq 0$ , wenn für die Wurzeln des Stabilitätspolynoms  $\pi(\lambda; qh)$  gilt:*

$$|\lambda_i| \leq 1 \quad \text{bzw.} \quad |\lambda_i| < 1 \quad \text{im Fall mehrfacher Wurzeln.}$$

Die Menge aller  $\bar{h} \in \mathbb{C}$ , für welche die LMM absolut stabil ist, wird wieder als ihr „Stabilitätsgebiet“ (kurz SG) bezeichnet. Die Konsistenz dieser Definition der absoluten Stabilität für LMMn mit der für Einschrittverfahren wird durch den folgenden Satz sichergestellt:



**Satz 4.5 (Lineare Differenzgleichungen):** Die Lösungen der homogenen linearen Differenzgleichung (4.3.18) bilden einen  $R$ -dimensionalen Vektorraum. Seien  $\lambda_1, \dots, \lambda_m$  ( $m \leq R$ ) die paarweise verschiedenen Wurzeln des Stabilitätspolynoms  $\pi(\lambda; qh)$  mit den Vielfachheiten  $\mu_1, \dots, \mu_m$ . Dann bilden (im Falle  $\alpha_0 - hq\beta_0 \neq 0$  und  $\alpha_R - hq\beta_R \neq 0$ ) die  $R = \sum_{i=1}^m \mu_i$  Folgen mit den Elementen

$$\begin{aligned} y_n &= \lambda_i^n \\ y_n &= n\lambda_i^n \\ &\vdots \\ &\quad (i = 1, \dots, m) \\ y_n &= n(n-1)\dots(n-\mu_i+2)\lambda_i^n \end{aligned} \tag{4.3.19}$$

eine Basis des Lösungsraumes.

**Beweis:** Zur Abkürzung setzen wir  $\gamma_r := \alpha_r - hq\beta_r$  und  $\pi(\lambda) := \pi(\lambda; qh)$ . Die Lösungsmenge der linearen homogenen Differenzgleichung

$$\sum_{r=0}^R \gamma_{R-r} y_{n-r} = 0$$

ist offensichtlich ein linearer Raum. Zu jedem Satz von Startwerten  $\{y_0, \dots, y_{R-1}\}$  ist eindeutig eine Lösung  $(y_n)_{n \in \mathbb{N}}$  bestimmt ( $\gamma_R \neq 0$ ):

$$y_n = -\frac{1}{\gamma_R} \sum_{r=1}^R \gamma_{R-r} y_{n-r}, \quad n \geq R.$$

Für ein System  $\{(y_n^{(i)})_{n \in \mathbb{N}}\}_{i=1, \dots, m}$  von Lösungen folgt daher aus

$$\sum_{i=1}^m c_i y_k^{(i)} = 0, \quad k=0, \dots, R-1,$$

notwendig auch

$$\sum_{i=1}^m c_i y_n^{(i)} = 0 \quad \forall n \geq 0.$$

Ein System  $\{(y_n^{(i)})_{n \in \mathbb{N}}\}_{i=1, \dots, m}$  von Lösungen ist also genau dann linear unabhängig, wenn es die Startwerte sind:

$$\sum_{i=1}^m c_i y_k^{(i)} = 0 \quad (k=0, \dots, R-1) \Rightarrow c_i = 0 \quad (i=1, \dots, m).$$

Folglich ist die Dimension des Lösungsraumes genau  $R$ . Wir zeigen nun zunächst, dass die obigen Folgen  $(y_n)_{n \in \mathbb{N}}$  Lösungen sind. Wegen  $\gamma_0 \neq 0$  ist  $\lambda = 0$  keine Wurzel von  $\pi(\lambda)$ . Sei nun  $\lambda$  eine Wurzel der Vielfachheit  $\mu$ ; diese ist dann auch  $\mu$ -fache Wurzel von  $p(\lambda) := \lambda^n \pi(\lambda)$ . Folglich gilt

$$p^{(j)}(\lambda) = 0, \quad 0 \leq j \leq \mu-1,$$



**Beweis:** Siehe G.Dahlquist: Stability and error bounds in the numerical integration of ordinary differential equations; Trans.Roy.Inst.Technol., Stockholm, Nr. 130 (1959), und P. Henrici [9; S. 275 f.]. Q.E.D.

**Beispiel 4.2:** (a) *Simpson-Regel:*

$$y_n - y_{n-2} = \frac{1}{3}h [f_n + 4f_{n-1} + f_{n-2}]$$

$$\rho(\lambda) = \lambda^2 - 1 \quad \sigma(\lambda) = \frac{1}{3}(\lambda^2 + 4\lambda + 1)$$

$$\pi(\lambda; \bar{h}) = (1 - \frac{1}{3}\bar{h})\lambda^2 - \frac{4}{3}\bar{h}\lambda - (1 + \frac{1}{3}\bar{h}).$$

Hier und besonders bei Methoden höherer Ordnung ist es kaum praktikabel,  $\lambda_i$  in Abhängigkeit von  $\bar{h}$  explizit anzugeben. Dagegen können gewisse Informationen durch Betrachtung von  $O(\bar{h})$ -Approximationen von  $\lambda_i$  gewonnen werden. Die beiden Wurzeln von  $\rho(\lambda)$  sind  $\lambda_1(0) = 1$  und  $\lambda_2(0) = -1$ . Wir machen also den Ansatz  $\lambda_1(\bar{h}) = 1 + \gamma_1\bar{h} + O(\bar{h}^2)$  und  $\lambda_2(\bar{h}) = -1 + \gamma_2\bar{h} + O(\bar{h}^2)$ . Nach Einsetzen in  $\pi(\lambda; \bar{h}) = 0$  erhalten wir durch Koeffizientenvergleich, dass notwendig  $\gamma_1 = 1$  und  $\gamma_2 = \frac{1}{3}$ , und folglich

$$\lambda_1 = 1 + \bar{h} + O(\bar{h}^2), \quad \lambda_2 = -1 + \frac{1}{3}\bar{h} + O(\bar{h}^2).$$

Für hinreichend kleines  $\bar{h}$  kann der  $O(\bar{h}^2)$ -Term vernachlässigt werden, und wir finden

$$\lambda_1 \sim 1 + \bar{h} < 1, \quad \lambda_2 \sim -1 + \frac{1}{3}\bar{h} < -1.$$

Für  $q \in \mathbb{R}$  und kleines (positives)  $h$  ist die Simpson-Methode also numerisch instabil. Dasselbe gilt, wie man leicht nachrechnet (Übungsaufgabe) auch für die Mittelpunktsregel.

(b) *Adams-Moulton-Formel ( $R = 3$ ):*

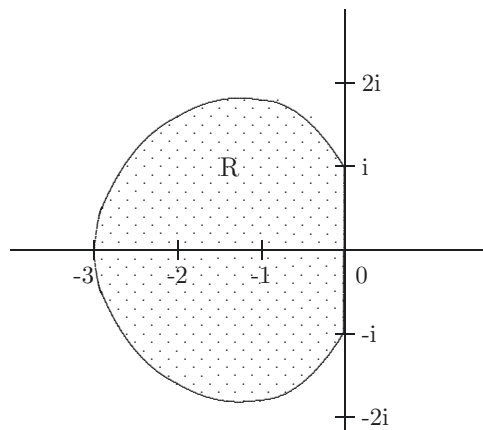


Abbildung 4.1: Stabilitätsgebiet der Adams-Moulton-Formel 3. Ordnung

(c) Stabilitätsintervalle  $(\alpha, 0)$  der Adams-Bashforth- und Adams-Moulton-Formeln:

Adams-Bashforth (explizit)				
$R$	1	2	3	4
$m$	1	2	3	4
$C_{m+1}$	$\frac{1}{2}$	$\frac{5}{12}$	$\frac{3}{8}$	$\frac{251}{720}$
$\alpha$	-2	-1	$-\frac{6}{11}$	$-\frac{3}{10}$

Adams-Moulton (implizit)				
$R$	1	2	3	4
$m$	2	3	4	5
$C_{m+1}$	$-\frac{1}{12}$	$-\frac{1}{24}$	$-\frac{19}{720}$	$-\frac{3}{160}$
$\alpha$	$-\infty$	-6	-3	$-\frac{90}{49}$

Zur Integration steifer AWA sollten Methoden verwendet werden, die möglichst große negative Werte von  $Re \bar{h}$  in ihrem Stabilitätsgebiet enthalten. Leider bedeutet die Forderung der A-Stabilität eine starke Einschränkung bei der Wahl der möglichen Methode.

**Satz 4.7 (A-stabile LMMn):** *Es gelten die folgenden Aussagen:*

1. Eine explizite LMM kann nicht A-stabil sein.
2. Die Ordnung einer A-stabilen impliziten LMM kann nicht größer als  $p = 2$  sein.
3. Die A-stabile implizite LMM der Ordnung  $p = 2$  mit kleinster Fehlerkonstante ist die Trapezregel  $y_n - y_{n-1} = \frac{1}{2}h(f_n + f_{n-1})$ .

**Beweis:** Siehe G.Dahlquist: A special stability problem for linear multistep methods. BIT 3, 27-43 (1963). Q.E.D.

Wegen der Ordnungsbegrenzung A-stabiler Methoden hat man folgenden abgeschwächten Stabilitätsbegriff eingeführt.

**Definition 4.6 (A( $\alpha$ )-Stabilität):** Eine LMM heißt „A( $\alpha$ )-stabil“,  $\alpha \in (0, \frac{\pi}{2})$ , wenn ihr Gebiet absoluter Stabilität den unendlichen Sektor  $\{-\alpha \leq \pi - \arg(\bar{h}) \leq \alpha\}$  enthält. Sie heißt „A(0)-stabil“, wenn ihr Stabilitätsintervall die ganze negative reelle Achse enthält.

Eine A(0)-stabile LMM ist geeignet zur Integration steifer AWA mit reellen Eigenwerten der Jacobi-Matrix  $f_x(t, u(t))$  (s. Übungsaufgabe).

**Satz 4.8 (A(0)-stabile LMM):** Eine explizite LMM kann nicht A(0)-stabil sein. Es existiert nur eine A(0)-stabile R-Schritt-Methode der Ordnung  $p \geq R + 1$ , nämlich die Trapezregel.

**Beweis:** Siehe O.B.Widlund: A note on unconditionally stable linear multistep methods. BIT 7, 65-70 (1967). Q.E.D.

**Beispiel 4.3:** Wir betrachten die allgemeine LMM

$$\sum_{r=0}^R \alpha_{R-r} y_{n-r} = h \sum_{r=0}^R \beta_{R-r} f_{n-r}.$$

A( $\alpha$ )-Stabilität erfordert, dass die Wurzeln des Stabilitätspolynoms  $\pi(\lambda; \bar{h})$  im Innern des Einheitskreises liegen für reelle  $\bar{h}$  mit  $\bar{h} \rightarrow -\infty$ . In diesem Limes gehen die Wurzeln von  $\pi(\lambda; \bar{h})$  über in die von  $\sigma(\lambda)$ . Daher ist es natürlich,  $\sigma(\lambda)$  so zu wählen, dass seine Wurzeln im Innern des Einheitskreises liegen, z. B.:  $\sigma(\lambda) = \beta_R \lambda^R$ . Wir erhalten so die uns schon bekannten Rückwärtsdifferenzenformeln (BDF)

$$\sum_{r=0}^R \alpha_{R-r} y_{n-r} = h \beta_R f_n. \quad (4.3.20)$$

Tabelle 4.2: Stabilitätsgebiete der Rückwärtsdifferenzenformeln

R	$\beta_R$	$\alpha_6$	$\alpha_5$	$\alpha_4$	$\alpha_3$	$\alpha_2$	$\alpha_1$	$\alpha_0$	$\alpha$
1	1						1	-1	$90^0$
2	$\frac{2}{3}$					1	$-\frac{4}{3}$	$\frac{1}{3}$	$90^0$
3	$\frac{6}{11}$				1	$-\frac{18}{11}$	$\frac{9}{11}$	$-\frac{2}{11}$	$88^0$
4	$\frac{12}{25}$			1	$-\frac{48}{25}$	$\frac{36}{25}$	$-\frac{16}{25}$	$\frac{3}{25}$	$73^0$
5	$\frac{60}{137}$		1	$-\frac{300}{137}$	$\frac{300}{137}$	$-\frac{200}{137}$	$\frac{75}{137}$	$-\frac{12}{137}$	$51^0$
6	$\frac{60}{147}$	1	$-\frac{360}{147}$	$\frac{450}{147}$	$-\frac{400}{147}$	$\frac{225}{147}$	$-\frac{72}{147}$	$\frac{10}{147}$	$18^0$

**Beispiel 4.4:** Die allgemeine 2-Schrittformel 2-ter Ordnung

$$y_n - (1 + \alpha) y_{n-1} + \alpha y_{n-2} = h \left\{ \left( \frac{1}{2} + \frac{\alpha}{2} + \beta \right) f_n + \left( \frac{1}{2} - \frac{3}{2}\alpha - 2\beta \right) f_{n-1} + \beta f_{n-2} \right\}$$

ist null-stabil für  $-1 \leq \alpha < 1$ . Sie ist A-stabil für  $\alpha > -1, \beta > -\frac{\alpha}{2}$ . Für  $\alpha = \frac{1}{3}, \beta = 0$  erhält man die *Rückwärtsdifferenzen-Formel*

$$y_n - \frac{4}{3} y_{n-1} + \frac{1}{3} y_{n-2} = \frac{2}{3} h f_n.$$

## 4.4 Praktische Aspekte

Wir wollen uns nun mit der praktischen Durchführung der Mehrschrittmethoden beschäftigen. Dazu gehört insbesondere wieder eine effektive Schrittweitenkontrolle.

#### 4.4.1 Berechnung von Startwerten

Die Startwerte  $y_0, \dots, y_{R-1}$ , kann man z. B. mit einem Einschrittverfahren (meistens vom Runge-Kutta-Typ) generieren. Nach  $R - 1$  Schritten hat man dabei gemäß (4.2.11) die Fehlerabschätzung (für  $y_0 = u(t_0)$ )

$$\max_{0 \leq \nu \leq R-1} \|y_\nu - u(t_\nu)\| \leq h \sum_{\nu=0}^{R-1} \|\tau_\nu^h\| e^{L R h}. \quad (4.4.21)$$

Ist das Einschrittverfahren von der Ordnung  $p^*$ , so werden die Startwerte  $y_0, \dots, y_{R-1}$ , mit der Ordnung  $p^* + 1$  berechnet. In Verbindung mit einer LMM der Ordnung  $p$  ist also zur Gewährleistung der globalen Fehlerordnung  $p$  eine Startprozedur der Ordnung  $p - 1$  erforderlich. Wegen des geringen Aufwandes für die  $R - 1$  Schritte dieser Startprozedur verwendet man in der Praxis jedoch meist Methoden der Ordnung  $p^* \geq p$ , um den Anfangsfehler klein zu halten.

Alternativ wird auch eine sog. „Selbststartprozedur“ verwendet, bei der die Startwerte mit Hilfe von Vertretern derselben Klasse von LMM aufsteigender Ordnung erzeugt werden. Bei der Lösung von potenziell steifen AWAn beginnt man z. B. innerhalb der Familie der Rückwärtsdifferenzenformeln im ersten Schritt mit der impliziten Euler-Formel 1. Ordnung und verwendet dann die so vorhandenen zwei Startwerte zur Anwendung der Rückwärtsdifferenzenformel 2. Ordnung und so weiter, bis ausreichend viele Startwerte erzeugt sind. Zur Erreichung einer ausreichenden Genauigkeit der Startprozedur werden die ersten Schritte niedriger Ordnung mit entsprechend kleineren Zeitschrittweiten durchgeführt. Dabei kommen sinnvollerweise LMM zu nicht äquidistanten Schrittweiten zum Einsatz.

#### 4.4.2 Lösung der impliziten Gleichungssysteme

Bei einer impliziten LMM

$$\sum_{r=0}^R \alpha_{R-r} y_{n-r} = h \sum_{r=0}^R \beta_{R-r} f_{n-r}, \quad \alpha_R = 1, \beta_R \neq 0,$$

ist bei berechneten Werten  $y_{n-R}, \dots, y_{n-1}$ , der neue Wert  $y_n$  bestimmt als Lösung des (i. Allg, nichtlinearen) Gleichungssystems

$$y_n - h\beta_R f(t_n, y_n) - g_n = 0 \quad (4.4.22)$$

wobei

$$g_n = h \sum_{r=1}^R \beta_{R-r} f_{n-r} - \sum_{r=1}^R \alpha_{R-r} y_{n-r}.$$

Wir haben schon im Stabilitätssatz 4.1 gesehen, dass dieses Gleichungssystem für

$$h < \frac{1}{L|\beta_R|} \quad (4.4.23)$$

( $L$  die Lipschitz-Konstante von  $f(t, x)$ ) eine eindeutig bestimmte Lösung besitzt. Diese kann man durch eine *Fixpunktiteration*

$$y_n^{(k+1)} = h\beta_R f(t_n, y_n^{(k)}) + g_n, \quad k = 0, 1, 2, \dots,$$

ausgehend vom einem beliebigen Startwert  $y_n^{(0)}$  bestimmen. Dabei gilt die Fehlerabschätzung

$$\|y_n^{(k)} - y_n\| \leq q^k \|y_n^{(0)} - y_n\| \quad (4.4.24)$$

mit  $q = hL|\beta_R| < 1$ . Im Fall moderater Lipschitz-Konstante  $L \sim 1$  bedeutet dies, dass sich die Genauigkeit der Näherung  $y_n^{(k)}$  in jedem Iterationsschritt um eine  $h$ -Potenz erhöht. Für eine LMM der Ordnung  $p$  gilt nach  $p$  Iterationsschritten:

$$\|y_n^{(p)} - u_n\| \leq \|y_n^{(p)} - y_n\| + \|y_n - u_n\| \leq ch^p. \quad (4.4.25)$$

Sinnvollerweise wird die Iteration also nach etwa  $p+1$  Schritten abgebrochen, da dann i. Allg. bereits das Genauigkeitsniveau des Diskretisierungsfehlers erreicht ist.

Die Bedingung (4.4.23) an die Schrittweite  $h$  bedeutet im Falle  $L \gg 1$  (*steifes Problem*) eine zu starke Einschränkung. Eine Alternative ist, wie schon bei den *impliziten* Einschrittverfahren diskutiert, die *Newton-Iteration*

$$\begin{aligned} [I - h\beta_R f_x(t_n, y_n^{(k)})] y_n^{(k+1)} &= [I - h\beta_R f_x(t_n, y_n^{(k)})] y_n^{(k)} - y_n^{(k)} + \\ &h\beta_R f(t_n, y_n^{(k)}) + g_n, \quad k = 0, 1, 2, \dots, \end{aligned}$$

welches z. B. für *monotone* AWAn garantiert konvergiert.

#### 4.4.3 Prädiktor-Korrektor-Methode

Eine in der Praxis wichtige Variante der Fixpunktiteration zur Lösung des Gleichungssystems (4.4.22) sind die sog. *Prädiktor-Korrektor-Methoden*. Da in die Fehlerabschätzung (4.4.24) der Anfangsfehler  $\|y_n^{(0)} - y_n\|$  eingeht, liegt es nahe, das Verfahren durch Wahl eines möglichst guten Startwertes zu verbessern. Das Prädiktor-Korrektor-Verfahren verschafft sich diesen Startwert durch einmalige Anwendung einer *expliziten* LMM, den sog. „Prädiktor“. Mit Hilfe einer *impliziten* LMM, dem sog. „Korrektor“, wird dieser Wert dann durch sukzessive Iteration korrigiert.

Wir wollen diese Methode an Hand einer häufig verwendeten Variante diskutieren:

*Prädiktor*: Adams-Bashforth-Formel 4-ter Ordnung

$$y_n = y_{n-1} + \frac{1}{24}h[55f_{n-1} - 59f_{n-2} + 37f_{n-3} - 9f_{n-4}], \quad (4.4.26)$$

*Korrektor*: Adams-Moulton-Formel 4-ter Ordnung

$$y_n = y_{n-1} + \frac{1}{24}h[9f_n + 19f_{n-1} - 5f_{n-2} + f_{n-3}]. \quad (4.4.27)$$

Diese Formeln haben die Fehlerkonstanten

$$C_5^{(P)} = \frac{251}{720}, \quad C_5^{(C)} = -\frac{19}{720}.$$

Die Prädiktor-Korrektor-Methode arbeitet dann wie folgt: Zu bereits berechneten Werten  $y_{n-r}$ ,  $r = 1, \dots, 4$ , wird durch Anwendung des Prädiktors „P“ der Startwert  $y_n^{(0)}$  bestimmt:

$$y_n^{(0)} = y_{n-1} + \frac{1}{24}h[55f_{n-1} - 59f_{n-2} + 37f_{n-3} - 9f_{n-4}],$$

bzw. in abstrakter Schreibweise

$$y_n^{(0)} = P(y_{n-1}, \dots, y_{n-4}).$$

Dann wird der Korrektor „C“ benutzt, um durch sukzessive Approximation den Wert  $y_n^{(0)}$  zu verbessern:

$$\begin{aligned} f_n^{(k-1)} &= f(t_n, y_n^{(k-1)}), \\ y_n^{(k)} &= y_{n-1} + \frac{1}{24}h[9f_n^{(k-1)} + 19f_{n-1} - 5f_{n-2} + f_{n-3}], \quad k = 1, 2, \dots, \end{aligned}$$

bzw. wieder in abstrakter Schreibweise

$$f_n^{(k-1)} = E(y_n^{(k-1)}), \quad y_n^{(k)} = C(y_{n-1}, \dots, y_{n-3}, f_n^{(k-1)}), \quad k = 1, 2, \dots$$

Wird die Iteration fortgeführt, bis eine vorgegebene Fehlerschranke  $\varepsilon$  erreicht ist, etwa

$$\|y_n^{(k)} - y_n^{(k-1)}\| \leq ch^4.$$

so spricht man von *Korrektur zur Konvergenz*. Es ist klar, dass in diesem Falle der schließlich akzeptierte Wert  $y_n := y_n^{(k)}$  im Wesentlichen unabhängig vom Startwert  $y_n^{(0)}$  ist, d. h.: Der Diskretisierungsfehler ist allein durch den Korrektor bestimmt.

Das *Korrigieren zur Konvergenz* ist in der Praxis meistens zu aufwendig, da in jedem Iterationsschritt eine weitere Funktionsauswertung  $f_n^{(k-1)} = E(y_n^{(k-1)})$  erforderlich ist. Daher begnügt man sich mit einer vorgegebenen Zahl  $k$  von Iterationen und akzeptiert  $y_n := y_n^{(k)}$  als neuen Wert. Das Verfahren wird dann bezeichnet als

$$P(EC)^k \quad \text{bzw.} \quad P(EC)^k E,$$

falls noch  $f_n^{(k)} = E(y_n^{(k)})$  berechnet wird. In der  $P(EC)^k$ -Form werden im Prädiktorschritt jeweils die Werte  $f_{n-r}^{(k-1)}$ ,  $r = 1, \dots, 4$ , verwendet. Für den Abschneidefehler dieser kombinierten Methode gilt dann folgende Aussage:

**Satz 4.9 (Ordnung des Prädiktor-Korrektor-Verfahrens):** Sei  $m^{(P)}$  die Ordnung des Prädiktors und  $m^{(C)}$  die des Korrektors. Dann ist die Ordnung  $m$  des Prädiktor-Korrektor-Verfahrens in  $P(EC)^k$ - oder  $P(EC)^k E$ -Form gegeben durch

$$m = \min\{m^{(C)}, m^{(P)} + k\}. \quad (4.4.28)$$

Im Falle  $m^{(C)} < m^{(P)} + k$  ist die Fehlerkonstante des kombinierten Verfahrens gleich der des Korrektors, d. h.:  $C_{m+1} = C_{m+1}^{(C)}$ .

**Beweis:** Übungsaufgabe.

Q.E.D.



#### 4.4.4 Fehlerschätzung und Schrittweitensteuerung: „Milne Device“

Die obige Aussage bzgl. der Fehlerkonstante des Prädiktor-Korrektor-Verfahrens lässt sich zur a posteriori Abschätzung des lokalen Abschneidefehlers und damit zur Schrittweitensteuerung verwenden. Diese Vorgehensweise hat Eingang in die Literatur gefunden als sog. „Milne Device“.

Für die mit den Formeln (4.4.26) und (4.4.27) gebildete kombinierte Methode gilt (unter der Annahme *exakter* Startwerte)

$$\begin{aligned} C_5^{(P)} h^5 u^{(5)}(t_n) &= u(t_n) - y_n^{(0)} + O(h^6) \\ C_5^{(C)} h^5 u^{(5)}(t_n) &= u(t_n) - y_n^{(k)} + O(h^6). \end{aligned}$$

Daraus erhalten wir offenbar

$$u^{(5)}(t_n) = \frac{y_n^{(k)} - y_n^{(0)}}{h^5(C_5^{(P)} - C_5^{(C)})} + O(h). \quad (4.4.29)$$

Im vorliegenden Fall ist der lokale Diskretisierungsfehler  $\hat{\tau}_n^h$  der kombinierten Methode von der Form

$$\hat{\tau}_n^h = C_5^{(C)} h^4 u^{(5)}(t_n) + O(h^5).$$

Mit Hilfe von (4.4.29) erhält man also eine Schätzung für  $\hat{\tau}_n^h(t_n)$  der Ordnung  $O(h^5)$  mit

$$\hat{\tau}_n^h = \frac{C_5^{(C)}}{C_5^{(P)} - C_5^{(C)}} \frac{y_n^{(k)} - y_n^{(0)}}{h} + O(h^5). \quad (4.4.30)$$

Mit dieser Schätzung für den Abschneidefehler  $\hat{\tau}_n^h$  lassen sich nun ähnlich wie bei den Einschrittverfahren Strategien zur Schrittweitenkontrolle angeben. Wir verzichten auf Angabe der Details.

Bei Vergrößerung der Schrittweite, etwa um den Faktor 2, werden nur bereits berechnete Werte  $y_{n-r}$  zusätzlich benötigt. Dies kann jedoch unter Umständen zu Speicherplatzproblemen führen. Generell lassen sich neue Startwerte bei Schrittweitenänderungen durch Einschrittverfahren beschaffen. Eine weitere Möglichkeit bei Schrittweitenverkleinerung zur Beschaffung der benötigten Zwischenwerte  $y_{n-r/2}$  besteht in der Interpolation der bereits berechneten Werte  $y_{n-r}$ . Dabei muss die Ordnung des Interpolationspolynoms natürlich der Ordnung der LMM angepasst sein.

**Numerischer Test:** Für die AWA

$$u'(t) = -200t u(t)^2, \quad t \geq -3, \quad u(-3) = \frac{1}{901},$$

mit der Lösung  $u(t) = (1 + 100t^2)^{-1}$  wurde der Wert  $u(0) = 1$  approximiert mit dem Prädiktor-Korrektor-Verfahren 4-ter Ordnung nach Adams-Bashforth-Moulton in PECE-Form mit dem Runge-Kutta-Verfahren 4-ter Ordnung als Startprozedur:

$$\begin{aligned} y_n^* &= y_{n-1} + \frac{1}{24}h \{55f_{n-1} - 59f_{n-2} + 37f_{n-3} - 9f_{n-4}\} \\ f_n^* &= f(t_n, y_n^*) \\ y_n &= y_{n-1} + \frac{1}{24}h \{9f_n^* + 19f_{n-1} - 5f_{n-2} + f_{n-3}\} \\ f_n &= f(t_n, y_n). \end{aligned}$$

Die Schrittweitensteuerung erfolgte gemäß *Milne Device*

$$\frac{C_5^{(C)}}{C_5^{(P)} - C_5^{(C)}} \frac{y_n - y_n^*}{h} \cong -0.07 \frac{y_n - y_n^*}{h} \sim \varepsilon = \text{eps} \frac{|y_{n-1}|}{h},$$

wobei erforderliche Zwischenwerte mit Hilfe der Runge-Kutta-Formel 4-ter Ordnung berechnet werden. Es zeigt sich, dass bei diesem einfachen Problem durch die adaptive Schrittweitenwahl keine Ersparnis hinsichtlich Rechenaufwand erreicht werden kann. Dafür kann aber ohne a priori Informationen die geforderte Genauigkeit garantiert werden.

Tabelle 4.3: Ergebnisse der adaptiven Rechnung bei 17-stelliger Arithmetik

Ordnung	eps	$h_{\min}$	$h_{\max}$	Fehler	# Auswertungen
$m = 4$	$10^{-17}$	$6 \cdot 10^{-5}$	$10^{-3}$	$2 \cdot 10^{-10}$	$\sim 16.500$
Rechnung mit konstanter (mittlerer) Schrittweite:					
		$4.5 \cdot 10^{-4}$		$2.7 \cdot 10^{-11}$	$\sim 13.400$

## 4.5 Übungsaufgaben

**Aufgabe 4.1:** Betrachtet werde die LMM

$$y_n + \alpha(y_{n-1} - y_{n-2}) - y_{n-3} = \frac{1}{2}(3 + \alpha)h(f_{n-1} + f_{n-2}),$$

mit einem freien Parameter  $\alpha \in \mathbb{R}$ .

- i) Man zeige, dass die angegebene Methode genau für  $\alpha \in (-3, 1)$  null-stabil ist.
- ii) Welche Konsistenzordnung hat das Verfahren für  $\alpha \in (-3, 1)$ ? Ist eine höhere Konsistenzordnung möglich?

**Aufgabe 4.2:** Zur Lösung der AWA

$$u''(t) = -20u'(t) - 19u(t), \quad t \geq 0, \quad u(0) = 1, \quad u'(0) = -10,$$

soll das Adams-Moulton-Verfahren dritter Ordnung

$$y_n = y_{n-1} + \frac{1}{12}h(5f_n + 8f_{n-1} - f_{n-2})$$

verwendet werden. Dazu forme man die Differentialgleichung zunächst in ein System erster Ordnung um. Wie klein muss dann die Schrittweite  $h$  bemessen sein, damit in jedem Zeitschritt die Konvergenz der Fixpunktiteration zur Berechnung von  $y_n$  garantiert ist?

**Aufgabe 4.3:** Man bestimme das Stabilitätsintervall und -Gebiet der expliziten Mittelpunktsregel:

$$y_n - y_{n-2} = 2hf_{n-1}$$

Man benutze hierzu den lokalen Störungsansatz zur Stabilitätsanalyse aus dem Text.

**Aufgabe 4.4 (Praktische Aufgabe):** Für das Modellproblem

$$u'(t) = -200t u(t)^2, \quad t \geq -3, \quad u(-3) = \frac{1}{901},$$

mit der Lösung  $u(t) = (1 + 100t^2)^{-1}$  soll näherungsweise der Wert  $u(0) = 1$  berechnet werden mit Hilfe:

- der klassischen 4-stufigen Runge-Kutta-Methode,
- der 4-Schritt-Adams-Bashforth-Methode,
- der 3-Schritt-Adams-Moulton-Methode.

Alle diese Formeln sind von 4-ter Ordnung. Als Startprozedur für die Mehrschrittverfahren werde die Runge-Kutta-Methode verwendet. Man führe die Rechnungen für die (äquidistanten) Schrittweiten  $h_i = 2^{-i}, i = 2, \dots, 8$ , durch und vergleiche die erreichte Genauigkeit und den erforderlichen Rechenaufwand (Anzahl der Funktionsauswertungen).

**Aufgabe 4.5:** Man zeige, dass die Rückwärtsdifferenzenformeln (BDF) der Stufen  $R = 1, 2, 3$ :

$$\sum_{r=0}^R \alpha_{R-r} y_{n-r} = h \beta_R f_n$$

konvergent sind. Welche A-stabilen Einschrittformeln wären (für  $R \geq 2$ ) jeweils als Startprozeduren geeignet?

**Aufgabe 4.6:** Man zeige durch Nachrechnen für ein Paar von einer expliziten und einer impliziten LMM:

$$\sum_{r=0}^R \alpha_{R-r}^{(P)} y_{n-r} = h \sum_{r=1}^R \beta_{R-r}^{(P)} f_{n-r}, \quad \sum_{r=0}^R \alpha_{R-r}^{(K)} y_{n-r} = h \sum_{r=0}^R \beta_{R-r}^{(K)} f_{n-r}.$$

a) Die Ordnung  $m$  des zugehörigen Prädiktor-Korrektor-Verfahrens in der P(EK)<sup>k</sup>E-Form ist  $m = \min\{m^{(K)}, m^{(P)} + k\}$ .

b) Gilt für die Ordnungen  $m^{(K)} < m^{(P)} + k$ , so ist die Fehlerkonstante  $C_{m+1}^*$  des Gesamtverfahrens gleich der Fehlerkonstante  $C_{m+1}^{(K)}$  des Korrektors.

**Aufgabe 4.7:** Zur Integration des steifen Systems

$$\begin{aligned}u'(t) &= -10u(t) - 100v(t), \\v'(t) &= 100u(t) - 10v(t), \\w'(t) &= u(t) + v(t) - tw(t),\end{aligned}$$

soll eine LMM möglichst hoher Ordnung verwendet werden. Welche von den im Text angegebenen Methoden sollte man nehmen?

**Aufgabe 4.8 (Praktische Aufgabe):** Man berechne eine Näherungslösung für die Modell-AWA

$$u'(t) = -200t u(t)^2, \quad t \geq -3, \quad u(-3) = \frac{1}{901},$$

mit Hilfe des Prädiktor-Korrektor-Verfahrens 4-ter Ordnung nach Adams-Bashforth-Moulton in PEKE-Form mit dem Runge-Kutta-Verfahren 4-ter Ordnung als Startprozedur:

$$\begin{aligned}y_n^* &= y_{n-1} + \frac{1}{24}h(55f_{n-1} - 59f_{n-2} + 37f_{n-3} - 9f_{n-4}), \quad f_n^* = f(t_n, y_n^*), \\y_n &= y_{n-1} + \frac{1}{24}h(9f_n^* + 19f_{n-1} - 5f_{n-2} + f_{n-3}), \quad f_n = f(t_n, y_n),\end{aligned}$$

mit zunächst festen Schrittweiten  $h = 0.2, 0.1, 0.05, 0.025, 0.0125, 0.00625$ .

a) Man vergleiche die berechneten Werte zum Zeitpunkt  $t = 3$  mit dem Wert  $u(3)$  der exakten Lösung  $u(t) = (1 + 100t^2)^{-1}$ .

b) In einem zweiten Schritt versuche man eine Schrittweitenstrategie auf der Basis des „Milne Device“

$$\left| \frac{C_5^{(K)}}{C_5^{(P)} - C_5^{(K)}} \frac{y_n - y_n^*}{h} \right| \approx -0.07 \left| \frac{y_n - y_n^*}{h} \right| \leq \varepsilon \frac{|y_{n-1}|}{h}$$

zu entwickeln. Dazu setze man  $\text{eps} = 10^{-16}$  und berechne die eventuell erforderlichen Zwischenwerte mit Hilfe der Runge-Kutta-Methode 4-ter Ordnung.

**Aufgabe 4.9:** Zur Rekapitulation der bisher im Text diskutierten „Stabilitätskonzepte“ beschreibe man kurz die Bedeutung der folgenden Begriffe und den Zusammenhang, in welchem sie verwendet werden:

- a) exponentielle Stabilität,
- b) diskrete Stabilität,
- c) numerische Stabilität,
- d) Null-Stabilität,
- e) A-Stabilität,
- f) A(0)-Stabilität.