

## 4 Lineare Mehrschrittverfahren

Als Verfahren höherer Ordnung haben wir bis jetzt RK-Verfahren kennengelernt. Bei RK-Verfahren steigt mit der gewünschten Ordnung des Verfahrens auch die benötigte Stufenzahl, d.h. die Anzahl Funktionsauswertungen: Für explizite RK-Verfahren werden für Ordnung  $p$  mindestens  $s = p$  Stufen benötigt (vgl. Satz 2.17) und für die implizite RK-Verfahren ebenfalls mindestens  $s = p/2$  ( $p = 2s$  wird für die Gaußverfahren erreicht—vgl. Abschnitt 3.1.1). Eine wichtige Motivation für die Entwicklung von Mehrschrittverfahren war, Verfahren hoher Ordnung zu erzeugen, die (zumindest bei expliziten Verfahren) pro Schritt nur *eine* Funktionsauswertung benötigen. Die grundlegende Beobachtung hierbei ist, daß bei einem Einschrittverfahren für die Berechnung von  $y_{i+1}$  *nur* der Wert  $y_i$  benötigt wird. Da bei Verfahren höherer Ordnung fast immer Taylorentwicklungen in der einen oder anderen Form im Hintergrund stehen, muß ein Einschrittverfahren durch *mehrere* Funktionsauswertungen “herausbekommen”, wie die gesuchte Lösung sich verhält.  $k$ -Schrittverfahren verwenden  $k$  Funktionswerte  $y_{i-j}$ ,  $j = 0, \dots, k-1$ , um den nächsten Funktionswert  $y_{i+1}$  zu bestimmen. Diese zusätzliche Information erlaubt es, sinnvolle “Extrapolationen” der Lösung zu machen und auf diese Weise effizientere Verfahren zu ermöglichen. Eine alternative Sichtweise ist, daß einmal durchgeführte Funktionsauswertungen mehrfach wiederverwendet werden.

### 4.1 Beispiele von linearen Mehrschrittverfahren

Lineare Mehrschrittverfahren (LMM = *linear multistep method*) werden meist für *konstante* Schrittweite  $h$  formuliert. In dem vorliegenden Kapitel nehmen wir deshalb an, daß die Schrittweite  $h$  konstant ist. Dies schließt Schrittweitensteuerung zunächst einmal aus; wir werden diesen Punkt kurz am Ende des Kapitels ansprechen. Die uniforme Schrittweite  $h$  impliziert insbesondere, daß die Zeitpunkte  $t_i$  von der Form

$$t_i = t_0 + ih \quad (4.1)$$

sind. Desweiteren erweist es sich als zweckmäßig, die Abkürzung

$$f_i = f(t_i, y_i), \quad i = 0, \dots, N, \quad (4.2)$$

einzuführen.

#### 4.1.1 Integrationsbasierte Verfahren: Adamsverfahren

Die Adamsverfahren<sup>1</sup> werden vor allem für *nicht-steife* Probleme eingesetzt. Sie werden mit 3 Parametern  $r, s, k \in \mathbb{N}_0$  beschrieben. Hier beschreibt  $k$  den Grad des verwendeten Interpolationspolynoms. Der Parameter  $s$  beschreibt, durch welche Punkte das Interpolationspolynom gelegt wird. Insbesondere wird sich herausstellen, daß für  $s = 0$  ein implizites Verfahren entsteht und für  $s \geq 1$  ein explizites.

Gegeben seien die Werte  $y_j$ ,  $j = 0, \dots, i$ . Gesucht ist der Wert  $y_{i+1}$ . Hierzu gehen wir von der Darstellung

$$y(t_{i+1}) = y(t_{i-r}) + \int_{t_{i-r}}^{t_{i+1}} f(t, y(t)) dt \quad (4.3)$$

---

<sup>1</sup>Adams, John Couch 1819–1892

aus. Um  $y(t_{i+1})$  sinnvoll zu approximieren, ersetzen die Funktion  $f(t, y(t))$  durch das Interpolationspolynom  $P_k$  vom Grad  $k$  durch die Punkte  $f(t_{i+1-s-j}, y(t_{i+1-s-j})) \approx f(t_{i+1-s-j}, y_{i+1-s-j})$  für  $j = 0, \dots, k$  und setzen dieses in das Integral ein. Es ergibt sich:

$$P_k(t) = \sum_{j=0}^k f_{i+1-s-j} L_j(t), \quad L_j(t) = \prod_{\substack{m=0 \\ m \neq j}}^k \frac{t - t_{i+1-s-m}}{t_{i+1-s-j} - t_{i+1-s-m}}$$

und damit, indem man  $y_{i+1}$  definiert als  $y_{i+1} = y_{i-r} + \int_{t_{i-r}}^{t_{i+1}} P_k(t) dt$

$$y_{i+1} = y_{i-r} + \sum_{j=0}^k f_{i+1-s-j} \int_{t_{i-r}}^{t_{i+1}} L_j(t) dt.$$

Das Integral kann wegen der Annahme (4.1), daß die Knoten äquidistant verteilt sind, besonders einfach bestimmt werden:

$$y_{i+1} = y_{i-r} + h \sum_{j=0}^k \beta_j f_{i+1-s-j}, \quad (4.4)$$

$$\beta_j = \frac{1}{h} \int_{t_{i-r}}^{t_{i+1}} L_j(t) dt = \int_{-r}^1 \prod_{\substack{m=0 \\ m \neq j}}^k \frac{x - (1 - s - m)}{(1 - s - j) - (1 - s - m)} dx.$$

Wie man sieht, sind die Parameter  $\beta_j$  unabhängig von  $f$  und  $h$  und können vorab bestimmt werden. Für  $s = 0$  liegt ein implizites Verfahren vor, weil für  $j = 0$  auch auf der rechten Seite von (4.4) der gesuchte Wert  $y_{i+1}$  über  $f_{i+1}$  auftritt. Für  $s \geq 1$  stehen auf der rechten Seite von (4.4) nur bereits bekannte Werte, so daß ein explizites Verfahren vorliegt.

**Beispiel 4.1** Wichtige Vertreter dieser Klasse von LMM sind

1. **Adams-Bashforth-Verfahren** (explizit): Hier ist  $s = 1, r = 0$ . Es ergibt sich dann z.B.

$$\begin{aligned} k=0 & \quad y_{i+1} = y_i + h f_i && \text{(explizites Eulerverfahren)} \\ k=1 & \quad y_{i+1} = y_i + h \frac{1}{2} (3f_i - f_{i-1}) \\ k=2 & \quad y_{i+1} = y_i + h \frac{1}{12} (23f_i - 16f_{i-1} + 5f_{i-2}) \\ k=3 & \quad y_{i+1} = y_i + h \frac{1}{24} (55f_i - 59f_{i-1} + 37f_{i-2} - 9f_{i-3}) \end{aligned}$$

2. **Adams-Moulton-Verfahren**<sup>2</sup> (implizit): Hier ist  $s = 0, r = 0$ . Es ergibt sich dann z.B.

$$\begin{aligned} k=0 & \quad y_{i+1} = y_i + h f_{i+1} && \text{(implizites Eulerverfahren)} \\ k=1 & \quad y_{i+1} = y_i + h \frac{1}{2} (f_{i+1} + f_i) && \text{(Trapezregel)} \\ k=2 & \quad y_{i+1} = y_i + h \frac{1}{12} (5f_{i+1} + 8f_i - f_{i-1}) \\ k=3 & \quad y_{i+1} = y_i + h \frac{1}{24} (9f_{i+1} + 19f_i - 5f_{i-1} + f_{i-2}) \end{aligned}$$

---

<sup>2</sup>Moulton, Forest Ray, 1872–1952 (US Astronom, bekanntester Schuler Edwin Hubble)

3. **Nyström-Verfahren** (explizit): Hier ist  $s = 1$ ,  $r = 1$ . Es ergibt sich dann z.B.

$$k = 0 \quad y_{i+1} = y_{i-1} + 2hf_i \quad (\text{Mittelpunktsregel})$$

4. **Milne-Simpson-Regeln** (implizit): Hier ist  $s = 0$ ,  $r = 1$ . Es ergibt sich dann z.B.

$$k = 2 \quad y_{i+1} = y_{i-1} + \frac{1}{3}h(f_{i+1} + 4f_i + f_{i-1}).$$

■

### 4.1.2 Verfahren basierend auf Differenzieren: BDF-Verfahren

Wir stellen nun das Konstruktionsprinzip von BDF-Verfahren (*backward differentiation formulas*) vor. BDF-Verfahren werden vor allem bei *steifen* Problemen eingesetzt.

Die Adams-Verfahren aus Abschnitt 4.1.1 basieren auf der Integralform (4.3) der exakten Lösung. Man kann Mehrschrittverfahren auch dadurch erhalten, daß man durch die  $k + 1$  Punkte  $(t_{i+1-j}, y_{i+1-j})$ ,  $j = 0, \dots, k$ , ein Polynom  $k$ -ten Grades legt und dann für den unbekannten Wert  $y_{i+1}$  fordert, daß die Differentialgleichung in  $t_{i+1}$  gilt, d.h. man definiert

$$P_k(t) = \sum_{j=0}^k y_{i+1-j} L_j(t), \quad L_j(t) = \prod_{\substack{m=0 \\ m \neq j}}^k \frac{t - t_{i+1-m}}{t_{i+1-j} - t_{i+1-m}}$$

und fordert

$$P'_k(t_{i+1}) = f(t_{i+1}, y_{i+1}).$$

Dies ist ein implizites Verfahren. Aufgrund der Annahme (4.1) kann dies wieder geschrieben werden in der Form

$$\sum_{j=0}^k \alpha_j y_{i+1-j} = hf_{i+1}, \quad \alpha_j = hL'_j(t_{i+1}), \quad (4.5)$$

wobei die Zahlen  $\alpha_j$  unabhängig sind von  $i$ . Es ergibt sich:

**Beispiel 4.2 (BDF-Formeln)** Die ersten Vertreter der Formel (4.5) sind gegeben durch

$$\begin{aligned} k = 1 \quad & y_{i+1} - y_i = hf_{i+1} \quad (\text{implizites Eulerverfahren}) \\ k = 2 \quad & y_{i+1} - \frac{4}{3}y_i + \frac{1}{3}y_{i-1} = h\frac{2}{3}f_{i+1} \\ k = 3 \quad & y_{i+1} - \frac{18}{11}y_i + \frac{9}{11}y_{i-1} - \frac{2}{11}y_{i-2} = h\frac{6}{11}f_{i+1}. \end{aligned}$$

Formal können BDF-Verfahren für jedes  $k \in \mathbb{N}$  erzeugt werden; allerdings werden sie für  $k \geq 7$  aus Stabilitätsgründen (siehe unten) nicht eingesetzt. ■

Wir haben mit den BDF-Verfahren eine Variante gezeigt, wie durch Interpolation und anschließende Differentiation ein numerisches Verfahren hergeleitet werden kann. Offensichtlich könnten weitere Varianten definiert werden oder auch Kombinationen aus dem bisher entwickelten Verfahren gebildet werden.

## Allgemeine Mehrschrittverfahren

Betrachtet man die Adamsverfahren in der Form (4.4) und die BDF-Verfahren in der Form (4.5), so erkennen wir, daß lineare Mehrschrittverfahren mit  $k$  Schritten durch eine allgemeine Beziehung der Form

$$\sum_{j=0}^k \alpha_{k-j} y_{i+1-j} = h \sum_{j=0}^k \beta_{k-j} f_{i+1-j} \quad (4.6)$$

für Parameter  $\alpha_j, \beta_j, j = 0, \dots, k$ , definiert werden können. Da diese Gleichung mit einer Konstanten (ungleich Null) durchmultipliziert werden kann, ohne sich zu verändern, können wir O.B.d.A. die *Normierungsbedingung*

$$\alpha_k = 1 \quad (4.7)$$

fordern. Um von einem *echten*  $k$ -Schrittverfahren sprechen zu können, werden wir weiterhin

$$|\alpha_0| + |\beta_0| \neq 0 \quad (4.8)$$

fordern. Das Verfahren (4.6) ist explizit, falls  $\beta_k = 0$ , andernfalls ist es implizit. Wir beobachten, daß die Annahme  $\alpha_k = 1$  für hinreichend kleine  $h$  und glatte Funktionen  $f$  die Gleichung (4.6) mit Hilfe des Banachschen Fixpunktsatzes für  $y_{i+1}$  gelöst werden kann.

**Lemma 4.3** *Sei  $G \subset \mathbb{R}^2$  offen und  $f \in C(G)$  lokal lipschitzstetig im zweiten Argument. Dann ist für gegebene Werte  $y_{i-k+1}, \dots, y_i$  die Bestimmungsgleichung (4.6) für hinreichend kleines  $h$  eindeutig lösbar.*

**Beweis:** Folgt aus dem Banachschen Fixpunktsatz. □

Ein  $k$ -Schrittverfahren benötigt zur Definition von  $y_{i+1}$  die Werte  $y_{i+1-j}, j = 1, \dots, k$ . Insbesondere werden die Werte  $y_j, j = 1, \dots, k-1$ , benötigt. In der Praxis verschafft man sich hinreichend gute Approximationen an die exakten Werte  $y(t_j), j = 1, \dots, k-1$ , indem man in einer Anlaufrechnung ein Einschrittverfahren (mit nötigenfalls kleiner Schrittweite) einsetzt, um die  $y_j, j = 1, \dots, k-1$ , zu erhalten.

## 4.2 Konsistenz von linearen Mehrschrittverfahren

Im Fall von (expliziten) Einschrittverfahren haben wir den Konsistenzfehler definiert als den Fehler, den das Verfahren in einem Schritt bei exakten Anfangsbedingungen macht, d.h.  $\tau(t_0, y_0, h) = y_{t_0, y_0}(t_0 + h) - (y_0 + h\Phi(t_0, y_0, h))$ . Für Einschrittverfahren haben wir außerdem in Bemerkung 2.8 gesehen, daß wir von einem Verfahren der Ordnung  $p$  nur sinnvoll sprechen können, wenn die Lösung  $y$  in  $C^{p+1}$  ist, da Fehlerordnungen typischerweise mit Taylorentwicklungen bestimmt werden. Dies bedeutet, daß bei Konsistenz der Ordnung  $p$  sinnvollerweise Funktionen  $f \in C^p$  betrachtet werden.

Bei  $k$ -Mehrschrittverfahren gehen wir analog vor. Um den Konsistenzfehler zu definieren, nimmt man an, daß eine Lösung  $y \in C^1(J)$  einer Differentialgleichung  $y' = f(t, y)$  gegeben sei und fragt nach dem Fehler, der in einem Schritt gemacht wird, wenn alle benötigten Werte exakt gegeben sind. Um den Begriff der Konsistenzordnung  $p \in \mathbb{N}$  zu definieren, wird man dann rechte Seiten  $f \in C^p$  betrachten.

**Definition 4.4 (Konsistenzfehler bei LMM)** Sei  $y \in C^1(J)$  für ein Intervall  $J$  Lösung der Differentialgleichung  $y' = f(t, y)$ . Sei ein LMM mit  $k$  Schritten der Form (4.6) gegeben. Dann bezeichnet für  $h > 0$  die Größe

$$\tau(t_0, y, h) = y(t_k) - y_k$$

den Konsistenzfehler der Lösung  $y$ , wobei  $y_k$  durch das LMM (4.6) gegeben ist, falls die  $y_i$ ,  $i = 0, \dots, k-1$ , als  $y_i = y(t_i)$ ,  $i = 0, \dots, k-1$ , gewählt sind. (Hier ist  $t_i = t_0 + ih$ , und es wird implizit angenommen, daß  $h$  hinreichend klein ist, um sicherzustellen, daß a) die Zeitpunkte  $t_i$ ,  $i = 0, \dots, k$ , in  $J$  sind und b) daß die Approximation  $y_k$  eindeutig bestimmt werden kann—vgl. Lemma 4.3.)

Wir sagen, daß das LMM (4.6) ein Verfahren der Ordnung  $p \in \mathbb{N}$  ist, falls für jede Funktion  $f \in C^p(G)$  (hier ist  $G \subset \mathbb{R}^2$  offen) folgendes gilt: Für jede kompakte Menge  $K \subset G$  existieren  $C > 0$  und  $\underline{h} > 0$ , so daß für all  $h \in (0, \underline{h})$  und alle  $(t_0, y_0) \in K$  gilt:

$$|\tau(t_0, y_{t_0, y_0}, h)| \leq Ch^{p+1}.$$

Hier ist  $y_{t_0, y_0}$  die Lösung des Anfangswertproblems (1.2). Ein Verfahren der Konsistenzordnung 1 heißt kurz konsistent.

**Bemerkung 4.5** Wie bei Einschrittverfahren ist man daran interessiert, einem Verfahren eine Konsistenzordnung zuzusprechen unabhängig von der konkret betrachteten rechten Seite  $f$ . Die Forderung  $f \in C^p(J)$  in der Definition ergibt sich aus der Beobachtung, daß man die Konsistenzordnung typischerweise durch Taylorentwicklung der Lösung (die dann in  $C^{p+1}$  ist) überprüft.

Im Kontext von LMM wird Konsistenz und Konsistenzordnung oft auch über den Abschneidefehler definiert wie in Bemerkung 4.7 erläutert. ■

In Bemerkung 3.8 beobachteten wir im Kontext von impliziten Verfahren, daß der Konsistenzfehler auch beschrieben werden kann durch (3.11), welches mißt, wie genau eine exakte Lösung der Differentialgleichung die Gleichung erfüllt, die das numerische Verfahren beschreibt. Analog können wir damit den Abschneidefehler  $R(t_0, y, h)$  an der Stelle  $t_0$

$$R(t_0, y, h) = \left( \sum_{j=0}^k \alpha_{k-j} y(t_{k-j}) - h \sum_{j=0}^k \beta_{k-j} y'(t_{k-j}) \right) \quad (4.9)$$

durch “Einsetzen einer exakten Lösung  $y \in C^1$  in das numerische Schema” definieren. Wie im Falle von impliziten Verfahren (vgl. Bemerkung 3.8) ergibt sich, daß der Konsistenzfehler  $\tau$  und der Abschneidefehler  $R$  vergleichbar sind. Der Einfachheit halber führen wir dies für global lipschitzstetige rechte Seiten  $f$  vor:

**Lemma 4.6** Sei  $G = J \times \mathbb{R}$  für ein Intervall  $J$ . Sei  $f \in C(G)$  und lipschitzstetig im zweiten Argument:

$$|f(t, z) - f(t, \hat{z})| \leq L|z - \hat{z}| \quad \forall (t, z), (t, \hat{z}) \in G.$$

Dann gilt für  $h$  mit  $0 < h < 1/(|\beta_k|L)$  und jede Lösung  $y \in C^1(J)$  der Differentialgleichung  $y' = f(t, y)$

$$(1 - h|\beta_k|L)|\tau(t, y, h)| \leq |R(t, y, h)| \leq (1 + h|\beta_k|L)|\tau(t, y, h)|.$$

(Hier müssen  $t$  und  $h$  natürlich so sein, daß das Intervall  $(t, t + kh) \subset J$ .)

**Beweis:** Der Beweis ist ähnlich zu dem Vorgehen im Beweis von Satz 3.7. Sei  $y \in C^1$  Lösung von  $y' = f(t, y)$ . Für die Definition des Konsistenzfehlers  $\tau(t, y, h)$  an der Stelle  $t = t_0$  müssen wir  $y_k$  mit  $y(t_k)$  vergleichen, wobei  $y_k$  implizit über (4.6) definiert ist mit  $y_j = y(t_j)$ ,  $j = 0, \dots, k-1$ . Es ergibt sich aus (4.6):

$$\sum_{j=1}^k [\alpha_{k-j} y(t_{k-j}) - h \beta_{k-j} f(t_{k-j}, y(t_{k-j}))] + \alpha_k y_k - h \beta_k f(t_k, y_k) = 0.$$

Andererseits gilt

$$R(t, y, h) = \left( \sum_{j=0}^k \alpha_{k-j} y(t_{k-j}) - h \sum_{j=0}^k \beta_{k-j} y'(t_{k-j}) \right).$$

Durch Subtraktion dieser beiden Gleichungen ergibt sich dann wegen  $f(t_{k-j}, y(t_{k-j})) = y'(t_{k-j})$

$$\alpha_k (y(t_k) - y_k) = R(t, y, h) - h \beta_k [f(t_k, y(t_k)) - f(t_k, y_k)].$$

Nach unserer Normierungskonvention (4.7) ist  $\alpha_k = 1$ , so daß sich mit der Lipschitzstetigkeit von  $f$  ergibt

$$(1 - h|\beta_k|L)|y(t_k) - y_k| \leq |R(t, y, h)| \leq (1 + h|\beta_k|L)|y(t_k) - y_k|.$$

Weil  $\tau(t, y, h) = |y(t_k) - y_k|$ , ergibt sich daraus die Behauptung.

Der Beweis zeigt, daß die Forderung einer globalen Lipschitzbedingung für  $f$  im zweiten Argument lokalisiert werden kann: man benötigt lediglich Lipschitzbedingungen im zweiten Argument in der Nähe des Punkte  $(t_k, y(t_k))$ .  $\square$

**Bemerkung 4.7 (alternative Definition von Konsistenzordnung)** Lemma 4.6 zeigt, daß Abschneidefehler und Konsistenzfehler vergleichbar sind. In der Literatur wird deshalb oft der Abschneidefehler als Konsistenzmaß bei LMM verwendet. Ein LMM wird dann ein Verfahren der Ordnung  $p$  genannt, wenn

$$|R(t, y, h)| = O(h^{p+1})$$

für alle  $y \in C^{p+1}$ . Genauer: Wenn es für jede Funktion  $y \in C^{p+1}(J)$  und jedes  $t \in J$  ein  $C(t) > 0$  gibt, so daß  $|R(t, y, h)| \leq C(t)h^{p+1}$  für hinreichend kleine  $h$ .

**Übung:** Nehmen Sie an, daß  $|R(t, y, h)| \leq C(t)h^{p+1}$ . Überlegen Sie sich, daß für feste kompakte Mengen  $K \subset J$  dann sogar gilt: Es existiert  $\underline{h} > 0$ , und es existiert  $C > 0$ , so daß für jedes  $h \in (0, \underline{h})$  und jedes  $t \in K$  gilt:  $|R(t, y, h)| \leq Ch^{p+1}$ .  $\blacksquare$

Für die Analyse eines LMM der Form (4.6) ist es zweckmäßig, die folgenden zwei Polynome einzuführen, welche auch *charakteristische Polynome* des LMM genannt werden:

$$\rho(\zeta) = \alpha_k \zeta^k + \alpha_{k-1} \zeta^{k-1} + \dots + \alpha_0 = \sum_{j=0}^k \alpha_j \zeta^j, \quad (4.10)$$

$$\sigma(\zeta) = \beta_k \zeta^k + \beta_{k-1} \zeta^{k-1} + \dots + \beta_0 = \sum_{j=0}^k \beta_j \zeta^j. \quad (4.11)$$

Mit Hilfe der Beobachtung aus Bemerkung 4.7 ist es einfach, die Konsistenzordnung für ein gegebenes LMM zu bestimmen: man überprüft mit Taylorentwicklung, ob für jede Funktion  $y \in C^{p+1}$  für den Abschneidefehler  $R(t, y, h) = O(h^{p+1})$  gilt. Dieses Überprüfen kann mit dem folgenden Satz vereinfacht werden:

**Satz 4.8** *Das lineare Mehrschrittverfahren (4.6) hat Konsistenzordnung  $p$  genau dann, wenn eine der folgenden vier äquivalenten Bedingungen erfüllt ist:*

- (i)  $R(t, \pi_p, h) = 0$  für alle Polynome  $\pi_p$  vom Grad  $p$ ;
- (ii)  $\sum_{j=0}^k \alpha_j = 0$  zusammen mit  $\sum_{j=0}^k \alpha_j j^q = q \sum_{j=0}^k \beta_j j^{q-1}$  für  $q = 1, \dots, p$ ;
- (iii)  $\rho(e^h) - h\sigma(e^h) = O(h^{p+1})$  für  $h \rightarrow 0$ ;
- (iv)  $\frac{\rho(\zeta)}{\ln \zeta} - \sigma(\zeta) = O(|\zeta - 1|^p)$  für  $\zeta \rightarrow 1$ ;

Insbesondere ist ein LMM konsistent (d.h.  $p \geq 1$ ), falls

$$\rho(1) = 0 \quad \text{und} \quad \rho'(1) = \sigma(1). \quad (4.12)$$

**Beweis:** Der Beweis wird in mehreren Schritten vollzogen.

1. *Schritt:* Wir zeigen zuerst, daß das Verfahren Konsistenzordnung  $p$  genau dann hat, wenn die Bedingungen aus (ii) erfüllt sind.

Entwickelt man die Funktionen  $y, y'$  in Taylorreihen um  $t_0$ , so ergibt sich

$$y(t_0 + ih) = \sum_{q=0}^p \frac{1}{q!} y^{(q)}(t_0) (ih)^q + O(h^{p+1}), \quad y'(t_0 + ih) = \sum_{q=0}^{p-1} \frac{1}{q!} y^{(q+1)}(t_0) (ih)^q + O(h^p). \quad (4.13)$$

Setzt man dies in die Definition von  $R(t, y, h)$  ein, so ergibt sich

$$\begin{aligned} R(t, y, h) &= \sum_{j=0}^k \left[ \alpha_j \sum_{q=0}^p \frac{j^q}{q!} h^q y^{(q)}(t_0) - h \beta_j \sum_{r=0}^{p-1} \frac{j^r}{r!} y^{(r+1)}(t_0) \right] + O(h^{p+1}) \\ &= y(t_0) \sum_{j=0}^k \alpha_j + \sum_{q=1}^p \frac{h^q}{q!} y^{(q)}(t_0) \left[ \sum_{j=0}^k \alpha_j j^q - q \sum_{j=0}^k \beta_j j^{q-1} \right] + O(h^{p+1}). \end{aligned}$$

Damit also  $R(t, y, h) = O(h^{p+1})$  für jedes (hinreichend glatte) Anfangswertproblem ist, müssen die Bedingungen aus (ii) gelten.

2. *Schritt:* Die Äquivalenz von (i) und (ii) folgt einfach: Für  $y \in \mathcal{P}_p$  gelten wie in (4.13) die Beziehungen

$$y(t_0 + ih) = \sum_{q=0}^p \frac{1}{q!} y^{(q)}(t_0) (ih)^q, \quad y'(t_0 + ih) = \sum_{q=0}^{p-1} \frac{1}{q!} y^{(q+1)}(t_0) (ih)^q.$$

Wiederholt man die Überlegungen aus dem 1. Schritt, so erkennt man, daß  $R(t, y, h) = 0$  genau dann, wenn die Bedingungen aus (ii) gelten.

3. *Schritt:* Wir zeigen nun die Äquivalenz von (ii) und (iii). Hierzu betrachten wir die Funktion  $R(t, \exp, h)$ , d.h. den Abschneidefehler für die spezielle Lösung  $y(t) = e^t$ . Offensichtlich ist

$$R(t, \exp, h) = \rho(e^h) - h\sigma(e^h).$$

Andererseits zeigt eine Rechnung, daß

$$R(t, \exp, h) = \sum_{j=0}^k \alpha_j + \sum_{q \geq 1} \frac{h^q}{q!} \left[ \sum_{j=0}^k \alpha_j j^q - q \sum_{j=0}^k \beta_j j^{q-1} \right].$$

Hieraus folgt nun die gewünschte Äquivalenz von (ii) und (iii).

4. *Schritt:* Mit der Transformation  $z = e^h$  (d.h.  $h = \ln z$ ) kann (iii) äquivalent geschrieben werden als

$$\rho(z) - \ln z \sigma(z) = O((\ln z)^{p+1}), \quad z \rightarrow 1. \quad (4.14)$$

Weil

$$\ln z = (z - 1) + O((z - 1)^2), \quad z \rightarrow 1,$$

ist damit (4.14) äquivalent mit (iv).

5. *Schritt:* Daß Konsistenz äquivalent ist mit den Bedingungen  $\rho(1) = 0$  zusammen mit  $\rho'(1) = \sigma(1)$  folgt unmittelbar aus (ii).  $\square$

Satz 4.8 erlaubt es, auf einfache Weise die Konsistenzordnung eines LMM zu überprüfen:

**Beispiel 4.9** Die Konsistenzordnung der Adams-Verfahren und der BDF-Verfahren kann mit Hilfe von Bedingung (i) von Satz 4.8 einfach bestimmt werden:

- Die Adams-Bashforth-Verfahren haben Konsistenzordnung  $k + 1$ : Setze  $p = k + 1$ . Bei Adams-Verfahren ist  $r = 0$  und die Verfahren sind von der Form

$$y_{i+1} - y_i = h \sum_{j=0}^k \beta_{i+1-j} f_{i+1-j}.$$

Die Koeffizienten  $\beta_j$  der Adams-Verfahren in (4.4) sind gerade so definiert, daß

$$\int_{t_{k-1}}^{t_k} P(t) dt = h \sum_{j=0}^k \beta_j P(t_j), \quad \forall P \in \mathcal{P}_k.$$

damit ergibt sich für jedes  $y \in \mathcal{P}_{k+1}$

$$y(t_k) - y(t_{k-1}) = \int_{t_{k-1}}^{t_k} y'(t) dt = h \sum_{j=0}^k \beta_j y'(t_j).$$

Also gilt für den Abschneidefehler  $R(t_0, y, h)$  für  $y \in \mathcal{P}_{k+1}$ , daß  $R(t_0, y, h) = 0$ , mithin gilt die Bedingung (i).

- Übung: Die BDF-Verfahren haben Konsistenzordnung  $k$ : Auch hier stellt man fest, daß für polynomielle Lösungen die Bedingung (i) erfüllt ist.

■



## 4.3 Stabilität und Konvergenz von Mehrschrittverfahren

### Einiges aus der Theorie der Differenzengleichungen

Wir betrachten die *homogene* Differenzengleichung der Form

$$a_k y_{n+k} + a_{k-1} y_{n+k-1} + \cdots + a_0 y_n = 0 \quad n \in \mathbb{N}_0, \quad (4.15)$$

d.h. wir betrachten auf dem Raum der Folgen  $(y_n)_{n \in \mathbb{N}} \subset \mathbb{C}$  die Menge aller Lösungen von (4.15). Wir nehmen ohne Beschränkung der Allgemeinheit an, daß  $a_k = 1$  und betrachten zudem nur den Fall  $a_0 \neq 0$ :

$$a_k = 1 \quad \text{und} \quad a_0 \neq 0. \quad (4.16)$$

Dann beobachten wir, daß die Menge der Lösungen von (4.15) ein  $k$ -dimensionaler Vektorraum ist:

- die Vektorraumeigenschaften (Linearität) folgen sofort aus (4.15);
- eine Lösung ist wegen  $a_k \neq 0$  festgelegt, sobald die  $k$  Startwerte  $y_0, \dots, y_{k-1}$  gegeben sind.

Der folgende Satz 4.10 gibt die Lösungen von (4.15) explizit an. Um die Aussagen des Satzes zu motivieren, definieren wir die Folge  $(Y_n)_{n=0}^\infty$  mit Elementen  $Y_n \in \mathbb{C}^k$ , welche gegeben ist durch

$$Y_n = \begin{pmatrix} y_n \\ y_{n+1} \\ \vdots \\ y_{n+k-1} \end{pmatrix}$$

Die Differenzengleichung (4.15) impliziert dann, daß für alle  $n \in \mathbb{N}_0$  gilt:

$$Y_{n+1} = \begin{pmatrix} y_{n+1} \\ y_{n+2} \\ \vdots \\ y_{n+k} \end{pmatrix} = AY_n = A \begin{pmatrix} y_n \\ y_{n+1} \\ \vdots \\ y_{n+k-1} \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 0 \\ -a_0 & -a_1 & \cdots & -a_{k-2} & -a_{k-1} \end{pmatrix}.$$

Damit ist  $Y_n$  gegeben durch

$$Y_n = AY_{n-1} = A^2 Y_{n-2} = \cdots = A^n Y_0 = X^{-1} J^n X Y_0,$$

wobei wir im letzten Schritt die Jordansche<sup>3</sup> Normalform  $A = X^{-1} J X$ , verwendet haben. Die Einträge von  $J$  werden bestimmt durch die Nullstellen des charakteristischen Polynoms  $\chi$  von  $A$ :

$$\chi(\lambda) = \det(A - \lambda \text{id}_k) = (-1)^k \sum_{j=0}^k a_j \lambda^j, \quad (4.17)$$

welches sich durch Entwicklung von  $\det(A - \lambda \text{id}_k)$  nach der letzten Zeile (Übung: man überzeuge sich davon!) ergibt. Diese Untersuchung zeigt, daß die Nullstellen des charakteristischen Polynoms  $\chi$  von entscheidender Bedeutung sind. Die Nullstellen zusammen mit ihrer Vielfachheit erlauben dann, die Jordansche Normalform zu bestimmen. Es ergibt sich insgesamt die folgende Darstellung für die Lösungen  $(y_n)_{n=0}^\infty$  der Differenzengleichung (4.15):

---

<sup>3</sup>Jordan, Camille, 1838–1922

**Satz 4.10** Sei (4.15) gegeben mit (4.16). Seien  $\lambda_i$ ,  $i = 1, \dots, m \leq k$  die paarweise verschiedenen Nullstellen des charakteristischen Polynoms  $\chi$ ; seien  $\mu_i$ ,  $i = 1, \dots, m$ , die Vielfachheiten der Nullstellen  $\lambda_i \in \mathbb{C}$ . Weiter seien  $k$  Folgen  $(y_n^{(i,j)})_{n=0}^\infty$ ,  $i = 1, \dots, m$  und  $j = 1, \dots, \mu_i$  wie folgt definiert: Für jedes  $i = 1, \dots, m$  betrachtet man die Folge

$$y_n^{(i,0)} = \lambda_i^n$$

sowie, falls die Vielfachheit  $\mu_i > 1$  ist, auch noch die  $\mu_i - 1$  Folgen

$$\begin{aligned} y_n^{(i,1)} &:= n\lambda_i^{n-1} \\ &\vdots \\ y_n^{(i,\mu_i-1)} &:= n(n-1)\cdots(n-\mu_i+2)\lambda_i^{n-\mu_i+1}. \end{aligned}$$

Diese  $k = \sum_{i=1}^m \mu_i$  Folgen bilden eine Basis des Lösungsraums der Differenzengleichung (4.15).

**Beweis:** Da wir bereits wissen, daß der Raum der Lösungen von (4.15) höchstens die Dimension  $k$  hat, genügt es zu zeigen, daß diese  $k$  Folgen Lösungen von (4.15) sind und zudem linear unabhängig.

1. *Schritt:* Wir zeigen, daß die angegebenen Folgen Lösungen von (4.15) sind. Wegen  $a_0 \neq 0$  ist damit  $\lambda = 0$  nicht Nullstelle von  $\chi$ , d.h.  $\lambda_i \neq 0$  für  $i = 1, \dots, m$ . Da  $\lambda_i$  eine  $\mu_i$ -fache Nullstelle von  $\chi$  ist, ist es auch für jedes  $n \in \mathbb{N}$  eine  $\mu_i$ -fache Nullstelle des Polynoms

$$\lambda \mapsto \tilde{\chi}(\lambda), \quad \tilde{\chi}(\lambda) = \lambda^n \chi(\lambda) = \sum_{j=0}^k a_j \lambda^{j+n}$$

d.h.

$$\tilde{\chi}^{(\mu)}(\lambda_i) = 0 \quad \mu = 0, \dots, \mu_i - 1.$$

Dies impliziert, daß für jedes  $n \in \mathbb{N}_0$

$$\begin{aligned} 0 &= \tilde{\chi}(\lambda_i) = \sum_{j=0}^k a_j \lambda_i^{n+j} = \sum_{j=0}^k a_j y_{n+j}^{(i,0)}, \\ &\vdots \\ 0 &= \tilde{\chi}^{(\mu_i-1)}(\lambda_i) = \sum_{j=0}^k a_j (j+n)(j+n-1)\cdots(j+n-\mu_i+2) \lambda_i^{j+n-\mu_i+1} = \sum_{j=0}^k a_j y_{n+j}^{(i,\mu_i-1)} \end{aligned}$$

Diese Aussagen bedeuten gerade, daß die für  $i = 1, \dots, m$  und  $j = 1, \dots, \mu_i$  oben definierten Folgen  $(y_n^{(i,j)})_{n=0}^\infty$  Lösungen der Differenzengleichung (4.15) sind.

2. *Schritt:* Zur Illustration der wesentlichen Beweisidee, daß die Folgen  $(y_n^{(i,j)})_{n \in \mathbb{N}_0}$  linear unabhängig sind, betrachten wir zunächst den Fall, daß die  $\lambda_i$  paarweise verschieden sind, d.h.  $m = k$  und  $\mu_i = 1$  für alle  $i$ .

Wir erinnern an die Eindeutigkeit der Polynominterpolation in den Punkten  $\lambda_i$ ,  $i = 1, \dots, k$ . D.h., die Matrix  $\mathbf{V} \in \mathbb{C}^{k \times k}$  gegeben durch

$$\mathbf{V} := \begin{pmatrix} \lambda_1^0 & \lambda_1^1 & \cdots & \lambda_1^{k-1} \\ \lambda_2^0 & \lambda_2^1 & \cdots & \lambda_2^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_k^0 & \lambda_k^1 & \cdots & \lambda_k^{k-1} \end{pmatrix}$$

ist regulär. Um die lineare Unabhängigkeit der Folgen  $(y_n^{(i,0)})_{n \in \mathbb{N}_0}$  zu zeigen, seien  $c_i \in \mathbb{C}$ ,  $i = 1, \dots, k$ , so daß

$$\sum_{i=1}^k c_i y_n^{(i,0)} = 0 \quad \forall n \in \mathbb{N}_0.$$

Wir zeigen nun, daß bereits

$$\sum_{i=1}^k c_i y_n^{(i,0)} = 0 \quad \forall n \in \{0, \dots, k-1\} \quad (4.18)$$

$c_1 = c_2 = \dots = c_k = 0$  implizieren. Aus (4.18) und der Darstellung der  $y_n^{(i,0)}$  folgt

$$\sum_{i=1}^k c_i \lambda_i^n = 0 \quad \forall n \in \{0, 1, \dots, k-1\}.$$

Für  $\mathbf{c} = (c_1, \dots, c_k)^\top$  und der oben definierten Matrix  $\mathbf{V}$  ist dies ein LGS von der Form

$$\mathbf{V}^\top \mathbf{c} = \mathbf{0}$$

Weil nach Obigem  $\mathbf{V}$  regulär ist, folgt auch die Invertierbarkeit von  $\mathbf{V}^\top$  und damit die gewünschte Aussage  $\mathbf{c} = \mathbf{0}$ .

*3. Schritt (Erinnerung an Hermiteinterpolation):* Genau wie im Fall der klassischen Polynominterpolation zeigt der Fundamentalsatz der Algebra, daß für die paarweise verschiedenen Punkte  $\lambda_i$ ,  $i = 1, \dots, m$ , und gegebenen  $f_{ij} \in \mathbb{C}$ ,  $i = 1, \dots, m$ ,  $j = 0, \dots, \mu_i - 1$ , die *Hermiteinterpolationsaufgabe*

$$\text{finde } f \in \mathcal{P}_{k-1} \text{ s.d. } f^{(j)}(\lambda_i) = f_{ij}, \quad \forall i = 1, \dots, m, \quad \forall j = 0, \dots, \mu_i - 1$$

eine *eindeutige* Lösung hat. Setzt man das Polynom  $f \in \mathcal{P}_{k-1}$  als  $f(z) = \sum_{n=0}^{k-1} f_n z^n$  an, so ist die Hermiteinterpolationsaufgabe ein LGS für die Koeffizienten  $f_n$ ,  $n = 0, \dots, k-1$ ; die Eindeutigkeitsaussage stellt sicher, daß die zugehörige Matrix  $\mathbf{V}$  regulär ist. Diese Matrix  $\mathbf{V} \in \mathbb{C}^{k \times k}$  hat die folgende Form

$$\mathbf{V}_{(i,j),n} = (z^n)^{(j)}|_{z=\lambda_i},$$

wobei wir den Spaltenindex  $n$  von 0 bis  $k-1$  laufen lassen und der Einfachheit halber für den Zeilenindex den Doppelindex  $(i,j)$  verwenden<sup>4</sup>. Die wesentliche Beobachtung ist nun, daß  $(z^n)^{(j)} = n(n-1) \dots (n-j+1) z^{n-j}$  und damit aus der Darstellung von  $y_n^{(i,j)}$  folgt:

$$\mathbf{V}_{(i,j),n} = y_n^{(i,j)}$$

*4. Schritt:* Wir zeigen, daß die  $k$  Folgen  $(y_n^{(i,j)})_{n=0}^\infty$  linear unabhängig sind. Seien  $c_{ij} \in \mathbb{C}$ , so daß

$$\sum_{i=1}^m \sum_{j=0}^{\mu_i-1} c_{ij} y_n^{(i,j)} = 0 \quad \forall n \in \mathbb{N}_0.$$

---

<sup>4</sup> $i = 1, \dots, m$ ,  $j = 0, \dots, \mu_i - 1$ ,  $\sum_{i=1}^m \mu_i = k$ , d.h. die Anzahl Zeilen von  $\mathbf{V}$  stimmt mit der Anzahl Spalten überein

Zu zeigen ist, daß daraus  $c_{ij} = 0$  für alle  $i, j$  folgt. Wir zeigen dies unter der schwächeren Voraussetzung

$$\sum_{i=1}^m \sum_{j=0}^{\mu_i-1} c_{ij} y_n^{(i,j)} = 0 \quad \forall n \in \{0, 1, \dots, k-1\}. \quad (4.19)$$

Wir erkennen nun, daß dieses  $k \times k$  LGS durch geeignete Numerierung der Unbekannten von der Form

$$\mathbf{V}^\top \mathbf{c} = \mathbf{0}$$

ist, wobei  $\mathbf{V} \in \mathbb{C}^{k \times k}$  im 3. Schritt definiert wurde. Weil die Matrix  $\mathbf{V}$  regulär ist, ist auch  $\mathbf{V}^\top$  regulär, was  $\mathbf{c} = \mathbf{0}$  nach sich zieht. Also ist  $c_{ij} = 0$  für  $i = 1, \dots, m$  und  $j = 0, \dots, \mu_i - 1$ .  $\square$

Von Interesse für uns werden homogene Differenzengleichungen sein, die *beschränkte* Lösungen für *beliebige* Startwerte  $y_0, \dots, y_{k-1}$  haben. Die geschlossene Darstellung der Lösungen der homogenen Differenzengleichung (4.15) in Satz 4.10 erlaubt uns eine einfache Charakterisierung der Differenzengleichungen mit dieser Eigenschaft:

**Lemma 4.11** *Folgende beiden Aussagen sind äquivalent:*

- (i) *Jede Lösung  $(y_n)_{n=0}^\infty$  von (4.15) ist (für  $n \rightarrow \infty$ ) beschränkt*
- (ii) *das Polynom  $\lambda \mapsto \chi(\lambda) := \sum_{j=0}^k a_j \lambda^j$  erfüllt das Wurzelkriterium, d.h., alle Nullstellen  $\lambda_i$  von  $\chi$  erfüllen  $|\lambda_i| \leq 1$  und Nullstellen  $\lambda_i$  mit  $|\lambda_i| = 1$  sind einfache Nullstellen.*

**Beweis:** Übung.  $\square$

### 4.3.1 Stabilität

Im Falle von Einschrittverfahren reichte die Konsistenz, um die Konvergenz des Verfahrens sicherzustellen. Bei LMM ist dies nicht mehr der Fall. Vielmehr muß eine weitere Bedingung, die sog. Stabilität, garantiert sein. Wir zeigen dies am folgenden Beispiel:

**Beispiel 4.12** Wir betrachten Zweischrittverfahren der Form (4.6). Es liegt nahe, die Koeffizienten  $\alpha_i, \beta_i$  so zu wählen, daß ein Verfahren möglichst hoher Konsistenzordnung entsteht. Eine Rechnung zeigt, daß ein Verfahren der Ordnung 3 erzeugt werden kann, welches die Form

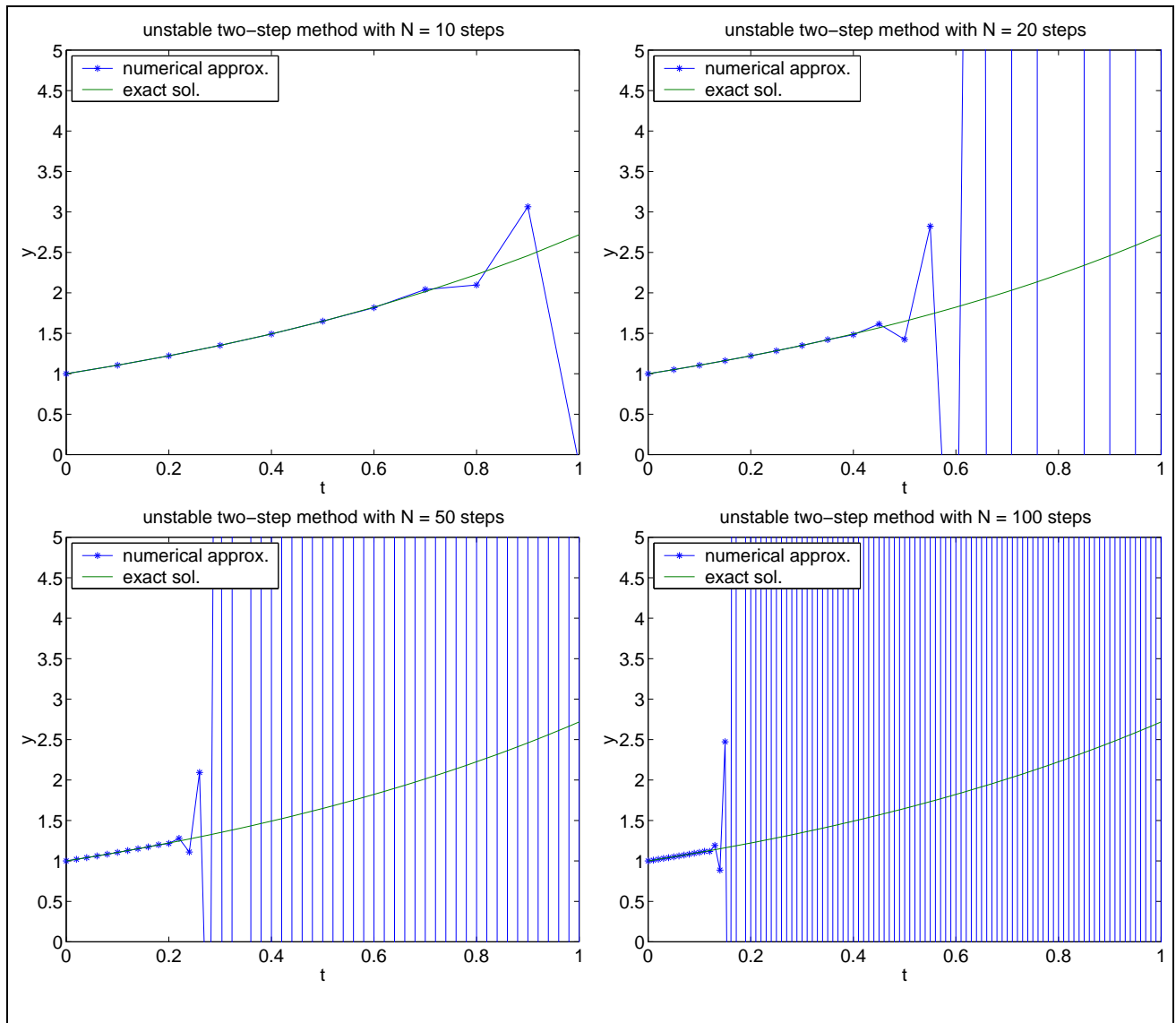
$$y_{i+1} + 4y_i - 5y_{i-1} = h(4f_i + 2f_{i-1}) \quad (4.20)$$

hat. Wir wenden nun dieses Verfahren auf das Anfangswertproblem  $y' = y$  mit  $y(0) = 1$  an. Wir verwenden die exakten Anfangsdaten  $y_0 = 1, y_1 = e^h$ . Die numerischen Ergebnisse für  $h = 1/10, h = 1/20, h = 1/50$  und  $h = 1/100$  sind in Fig. 4.1 wiedergegeben. Maximale Funktionswerte sind

h	1/10	1/20	1/50	1/100
$y_{max}$	3.06	$1.6 \cdot 10^6$	$3.7 \cdot 10^{25}$	$2.0 \cdot 10^{59}$

Um das Versagen zu verstehen, bestimmen wir die Lösung des numerischen Verfahrens. Das numerische Verfahren liefert Werte  $y_i$ , welche die (homogene) Differenzengleichung

$$y_{i+1} + 4(1-h)y_i - (5+2h)y_{i-1} = 0 \quad i = 1, \dots \quad (4.21)$$



**Figur 4.1:** Verhalten des instabilen Zweischrittverfahrens aus Beispiel 4.12.

Der Lösungsraum ist ein Vektorraum (Übung: man überzeuge sich davon) und seine Dimension ist zwei (die Folge  $(y_i)_{i=0}^\infty$  ist festgelegt, sobald  $y_0$  und  $y_1$  bestimmt sind). Eine explizite Darstellung der Lösungen von (4.21) erhält man, indem man den *Ansatz*  $y_i = \lambda^i$  für zu bestimmendes  $\lambda$  macht (siehe unten für eine genauere Betrachtung von Differenzengleichungen). Setzt man diesen Ansatz in (4.21) ein, so ergibt sich, daß  $\lambda$  die quadratische Gleichung

$$\lambda^2 + 4(1-h)\lambda - (5+2h) = 0$$

erfüllen muß. Eine Rechnung zeigt dann, daß

$$\lambda_1(h) = 1 + h + O(h^2), \quad \lambda_2 = -5 + O(h).$$

Damit haben die Lösungen der Differenzengleichung (4.21) die Form

$$y_i = A(\lambda_1(h))^i + B(\lambda_2(h))^i, \quad i = 0, 1, \dots,$$

mit Parametern  $A, B \in \mathbb{R}$ . Wir erkennen nun den Grund des Versagens des Beispiels: Ist  $B \neq 0$ , so ist die Approximation an der festen Stelle  $t_i > 0$  von der Form

$$y_i = A(\lambda_1(h))^{t_i/h} + B(\lambda_2(h))^{t_i/h}.$$

Damit die Approximation  $y_i$  beschränkt bleibt (eine Voraussetzung für Konvergenz!), muß wegen  $|\lambda_2(h)| > 1$  also  $B = 0$  oder zumindest sehr klein sein. Dies ist aber hier nicht gegeben. Vielmehr machen wir folgende Beobachtungen:

- Es kann sein, daß die Wahl der Anfangsbedingungen  $y_0$  und  $y_1$  erzwingt, daß  $B \neq 0$ . Man beachte, daß selbst "kleine" Werte von  $B$ , d.h.  $B = O(h^\gamma)$  für ein  $\gamma > 0$  nicht ausreicht, die Beschränktheit von  $y_i$  für ein festes  $t_i$  zu garantieren.
- Selbst wenn die Anfangsdaten  $y_0, y_1$  so sind, daß  $B = 0$ , so werden unvermeidliche Rundungsfehler dazu führen, daß nach einigen Schritten eine Lösung der Differenzengleichung erhalten wird, die zu Anfangsdaten gehört, bei denen  $B \neq 0$  ist.

Die Lösungen, die zur Nullstelle  $\lambda_2(h)$  gehören, sind offensichtlich unerwünscht und werden in der Literatur als *parasitäre* Lösungen bezeichnet. ■

Betrachtet man ein allgemeines LMM angewandt auf die Differentialgleichung  $y' = y$ , erhält man die Differenzengleichung

$$\sum_{j=0}^k (\alpha_{k-j} - h\beta_{k-j})y_{i+1-j} = 0. \quad (4.22)$$

Für  $h \rightarrow 0$  (wir sind ja an kleinen  $h$  interessiert) ergibt sich damit, daß wir die Lösungen der Differenzengleichung

$$\sum_{j=0}^k \alpha_{k-j}y_{i+1-j} = 0$$

betrachten sollten. Die Lösungen einer solchen homogenen Differenzgleichung hatten wir in Satz 4.10 und Lemma 4.11 diskutiert. Wesentlich ist das Verhalten der Nullstellen des Polynoms

$$\lambda \mapsto \rho(\lambda) = \sum_{j=0}^k \alpha_j \lambda^j,$$

welches wir als das erste charakteristische Polynom erkennen. Um Konvergenz zu erhalten, müssen wir mindestens Beschränktheit (für feste Stelle  $T$  und  $h \rightarrow 0$ ) der Approximationen fordern. Um parasitäre Lösungskomponenten zu unterdrücken, müssen wir also in Hinblick auf Lemma 4.11 fordern, daß  $\rho$  das Lemma 4.11 definierte Wurzelkriterium erfüllt. Bei LMM ist diese Bedingung allgemein als *Nullstabilität* bekannt:

**Definition 4.13 ((Null-)Stabilität von LMM)** *Ein LMM heißt nullstabil (auch kurz: stabil), falls das erste charakteristische Polynom  $\rho$  die folgende Wurzelbedingung erfüllt: Die Nullstellen  $\lambda$  von  $\rho$  erfüllt  $|\lambda| \leq 1$ , und die Nullstellen  $\lambda$  mit  $|\lambda| = 1$  sind einfache Nullstellen.*

**Bemerkung 4.14** (i) Eine interessante Interpretation der Nullstabilitätsbedingung ist, daß die zugehörige Differenzengleichung die exakte Lösung für die Differentialgleichung  $y' = 0$ , d.h.  $f \equiv 0$  ist. Somit fordert Nullstabilität, daß das LMM angewendet auf diese triviale Differentialgleichung beschränkte Lösungen liefert, selbst wenn die Anfangsdaten  $y_0, \dots, y_{k-1}$  beliebig gewählt sind. Wir werden diesen Punkt in Satz 4.22 noch einmal bearbeiten.

(ii) Man kann die Betrachtung des Polynoms  $\rho$ , d.h. die Reduktion des Problems (4.22) auf den Fall  $h = 0$  auch so sehen, daß, wegen der stetigen Abhängigkeit der Nullstellen  $\lambda(h)$  von  $h$ , die Forderung  $|\lambda(h)| \leq 1 + O(h)$  für  $h > 0$  die Forderung  $|\lambda(0)| \leq 1$  impliziert.

(iii) Der Notwendigkeit der Nullstabilität wurde oben anhand von Beispielen plausibel gemacht. Die mathematische Notwendigkeit der Nullstabilität ergibt sich durch Studium des Kerns des Konvergenzbeweises für LMM (siehe Satz 4.20 unten). Dort geht wesentlich ein, daß man für nullstabilen  $k$ -Schritt-Verfahren eine Matrixnorm  $\|\cdot\|$  finden kann, so daß die zugehörige "Begleitmatrix"  $A$  nach Satz 4.19 die Bedingung  $\|A\| = 1$  erfüllt—siehe die Fußnote auf S. 75.

■

**Beispiel 4.15** Die Verfahren aus Beispiel 4.1 sind nullstabil. Man überlegt sich leicht, daß für die expliziten und impliziten Adamsverfahren das zugehörige erste charakteristische Polynom die Form  $z^{k+1} - z^k$  hat, dessen Nullstellen die einfache Nullstelle  $z = 1$  und die  $k$ -fache Nullstelle  $z = 0$ . Ebenso beim Nyström-Verfahren und bei der Milne-Simpson Regel, bei denen  $\rho(z) = z^{k+1} - z^{k-1}$ .

■

Die Stabilität der BDF-Verfahren ist etwas komplizierter zu analysieren:

**Satz 4.16 (Stabilität der BDF-Verfahren)** *Die  $k$ -Schritt BDF-Verfahren aus Beispiel 4.2 sind nullstabil für  $k \leq 6$ . Sie sind nicht nullstabil für  $k \geq 7$ .*

**Beweis:** Wir verweisen auf [6, Chap. III.3, Thm. 3.6] für den Beweis. Der Beweis der Stabilität für  $k \leq 6$  erfolgt durch direktes Nachrechnen. Funktionentheoretische Argumente zeigen, daß kein BDF-Verfahren für  $k \geq 7$  nullstabil sein kann.  $\square$

Betrachtet man LMM der Form (4.6), so sind  $2k + 1$  Koeffizienten frei wählbar. Man kann deshalb hoffen, daß man LMM der Ordnung  $2k$  konstruieren kann. Die Forderung nach Nullstabilität eines LMM stellt zusätzliche Bedingungen an die Koeffizienten  $\alpha_i$  und wird somit die maximal erreichbare Ordnung reduzieren. In der Tat zeigt der folgende Satz, daß die Forderung nach Nullstabilität die maximal erreichbare Ordnung auf ungefähr die Hälfte reduziert:

**Satz 4.17 (erste Dahlquist Schranke)** *Die maximal erreichbare Ordnung  $p$  eines nullstabilen LMM erfüllt die Bedingungen*

$$\begin{aligned} p &\leq k + 2 && \text{falls } k \text{ gerade ist,} \\ p &\leq k + 1 && \text{falls } k \text{ ungerade ist,} \\ p &\leq k && \text{falls } \beta_k/\alpha_k \leq 0. \end{aligned}$$

*Insbesondere ist damit für explizite Verfahren (dort ist  $\beta_k = 0!$ ) die maximal erreichbare Ordnung  $k$ .*

**Beweis:** Wir verweisen auf den Beweis von [6, Chap. III.3, Thm. 3.5]. □

Nullstabile LMM, die diese Dahlquistschranken erreichen, heißen *optimale* Mehrschrittverfahren.

### 4.3.2 Konvergenz

Genau wie bei Einschrittverfahren wollen wir von Konvergenz sprechen, wenn die numerischen Approximation  $y_i$  gegen den exakten Wert  $y(t_i)$  konvergieren:

$$\max_{i=0,\dots,N} |y_i - y(t_i)| \rightarrow 0.$$

Damit dies erfüllt ist, müssen die Anfangswerte  $y_0, \dots, y_{k-1}$ , die in der Praxis durch ein weiteres numerisches Verfahren (z.B. ein Einschrittverfahren) erhalten werden, ebenfalls konvergieren. Wegen der zusätzlich eingeführten Fehler durch die Anfangsdaten, ist unsere formale Definition von Konvergenz wie folgt:

**Definition 4.18 (Konvergenz von LMM)** *Wir sagen, daß ein  $k$ -Schrittverfahren angewandt auf ein Anfangswertproblem  $y' = f(t, y)$ ,  $y(t_0) = y_0$  konvergent auf  $[t_0, T]$  ist, falls*

$$\max_{i=0,\dots,N} |y_i - y(t_i)| \rightarrow 0 \quad \text{für } h \rightarrow 0$$

*gilt, wenn nur die Anfangsdaten die Bedingung*

$$\max_{i=0,\dots,k-1} |y_i - y(t_i)| \rightarrow 0 \quad \text{für } h \rightarrow 0$$

*erfüllen.*

Wir haben in mit den vorangegangenen Überlegungen Konsistenz und Nullstabilität als Voraussetzungen für Konvergenz herausgearbeitet (Satz 4.22 formuliert dies präzise mit dem Konvergenzbegriff aus Definition 4.18). Diese beiden Bedingungen erweisen sich bereits als hinreichend wie wir im Satz 4.21 unten zeigen werden. Um Satz 4.21 zu zeigen, benötigen wir zwei Ergebnisse als Vorbereitung.

**Satz 4.19** *Zu jeder Matrix  $A \in \mathbb{R}^{k \times k}$  gibt es für jedes  $\varepsilon > 0$  eine Norm  $\|\cdot\|$  auf dem  $\mathbb{R}^k$ , so daß die zugeordnete Matrixnorm  $\|\cdot\|$  die Abschätzung*

$$\|A\| \leq \rho(A) + \varepsilon,$$



wobei  $\rho(A) = \max\{|\lambda| \mid \lambda \in \mathbb{C} \text{ ist Eigenwert von } A\}$  der Spektralradius von  $A$  ist. Ist zudem jeder Eigenwert  $\lambda$  mit  $|\lambda| = \rho(A)$  ein einfacher Eigenwert, dann kann die Matrixnorm  $\|\cdot\|$  so gewählt werden, daß

$$\|A\| = \rho(A).$$

**Beweis:** Sei  $J = TAT^{-1}$  die Jordansche Normalform von  $A$ , d.h. die Matrix  $J$  hat die Form

$$J = \begin{pmatrix} J_1 & & \\ & J_2 & \\ & & \ddots \\ & & & J_m \end{pmatrix}, \quad J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix} \in \mathbb{C}^{k_i \times k_i}$$

wobei für die Dimensionen der  $m$  Jordanblöcke  $J_m$  gilt:  $\sum_{i=1}^m k_i = k$ . Wir beachten, daß

$$\rho(A) = \max_{i=1, \dots, m} |\lambda_i|.$$

Wir definieren nun für  $\varepsilon > 0$  die Diagonalmatrix  $D_\varepsilon \in \mathbb{R}^{k \times k}$  durch  $D_\varepsilon = \text{diag}(1, \varepsilon, \varepsilon^2, \dots, \varepsilon^{k-1})$ . Dann ist<sup>5</sup>

$$D_\varepsilon^{-1} J D_\varepsilon = J^\varepsilon = \begin{pmatrix} J_1^\varepsilon & & \\ & J_2^\varepsilon & \\ & & \ddots \\ & & & J_m^\varepsilon \end{pmatrix}, \quad J_i^\varepsilon = \begin{pmatrix} \lambda_i & \varepsilon & & \\ & \lambda_i & \varepsilon & \\ & & \ddots & \ddots \\ & & & \ddots & \varepsilon \\ & & & & \lambda_i \end{pmatrix} \in \mathbb{C}^{k_i \times k_i}.$$

Dann gilt für die  $\|\cdot\|_\infty$ -Norm<sup>6</sup>

$$\|D_\varepsilon^{-1} T A T^{-1} D_\varepsilon\|_\infty = \|D_\varepsilon^{-1} J D_\varepsilon\|_\infty \leq \rho(A) + \varepsilon.$$

Da für eine beliebige nichtsinguläre Matrix  $B \in \mathbb{C}^{k \times k}$  der Ausdruck  $\|x\| := \|Bx\|_\infty$  eine Norm auf  $\mathbb{R}^k$  definiert und die zugehörige Matrixnorm durch  $\|M\| = \|BMB^{-1}\|_\infty$  ( $M \in \mathbb{R}^{k \times k}$ ) definiert wird, ist die gewünschte Norm durch  $\|x\| := \|D_\varepsilon^{-1} T x\|_\infty$  konstruierbar.

Wir betrachten nun den Fall, daß alle Eigenwerte  $\lambda$  mit  $|\lambda| = \rho(A)$  Vielfachheit 1 haben. Seien o.B.d.A. die Eigenwerte nach Größe absteigend sortiert, d.h.  $|\lambda_1| = |\lambda_2| = \dots = |\lambda_s| > |\lambda_{s+1}| \geq \dots \geq |\lambda_k|$ . Dann hat die Jordansche Normalform die Form

$$J = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_s & \\ & & & J_{s+1} \\ & & & & \ddots \\ & & & & & J_m \end{pmatrix}$$

Wähle nun  $\varepsilon = \rho(A) - |\lambda_{s+1}|$  (der Spezialfall  $s = k$  geht analog). Wir überprüfen

$$\|D_\varepsilon^{-1} T A T^{-1} D_\varepsilon\|_\infty = \|D_\varepsilon^{-1} J D_\varepsilon\|_\infty = \rho(A).$$

Somit leistet die gleiche Vektornorm wie oben das gewünschte.  $\square$

Wesentliches Hilfsmittel für unsere Konvergenzresultat, Satz 4.21, ist eine Analyse von gestörten Differenzengleichungen:

<sup>5</sup>NR: für beliebige Matrix  $J$  und  $(D_\varepsilon)_{ij} = \delta_{ij} \varepsilon^{i-1}$  ist  $(D_\varepsilon^{-1} J D_\varepsilon)_{ij} = \sum_k \sum_l \delta_{ik} \varepsilon^{-(i-1)} J_{kl} \delta_{lj} \varepsilon^{j-1} = \varepsilon^{-(i-1)} J_{ij} \varepsilon^{j-i} = J_{ij} \varepsilon^{j-i}$ . Wenn  $J$  Jordanform hat, folgt dann die angeg. Form.

<sup>6</sup>zur Erinnerung:  $\|B\|_\infty = \max_i \sum_j |B_{ij}|$