

Numerische Mathematik – Vorlesungs-Script

Prof. Martin Gutknecht

19. Oktober 2008

Teil I. Numerische Methoden für Anfangswertprobleme von gewöhnlichen Differentialgleichungen

Rohfassung:	Bearbeitung:	Korrekturen:
Andrin Schmidt	Andreas Steiger	Philipp Arbenz

Hinweis: Dies ist eine Version des Skripts, die unserer Vorstellung von “fertig” sehr nahe kommt. Alle Abschnitte wurden mehrfach korrigiert und das Skript ist bis auf kleine Details vollständig. Trotzdem sind wir um weitere Korrekturen froh.

Changelog:

- 08. 07. 2006:
Erste finale Version.
- 28. 08. 2006:
 - Beweis zum Hilfssatz in Kapitel 4.7 verdeutlicht.
 - Nummerierungen im ersten Kapitel hinzugefügt (Danke an Sibylle Arnold).
 - 2. Grafik in Kapitel 1.1 korrigiert.
 - Butcher-Tableau: Bemerkung zu expliziten Verfahren hinzugefügt.
 - Viele kleine Fehler verbessert.
- 08. 09. 2006:
Viele kleine Fehler verbessert; Insbesondere heisst das Charakteristische Polynom jetzt korrekterweise ρ statt δ .

Bekannte Fehler:

- In (3.14) ist die obere Klammer (von $h\hat{\Phi}$) zu lang – Sie müsste vor dem ersten f aufhören.

Warnung: Wir sind sicher dass diese Notizen eine Menge Fehler enthalten. Betreten der Baustelle auf eigene Verantwortung! Falls ihr einen entdeckt, schreibt eine Mail an mitschriften@vmp.ethz.ch, wir werden uns dann darum kümmern. Bitte erwähnt immer von welcher Version (die Id Zeile unten) ihr ausgeht. Weitere Informationen gibts unter:

<http://vmp.ethz.ch/wiki/index.php/Vorlesungsmitschriften>

\$Id\$

Inhaltsverzeichnis

1	Problemstellung, Notation	1
1.1	1 Differentialgleichung 1. Ordnung	1
1.2	Systeme von m Differentialgleichungen 1. Ordnung	1
1.3	Vektornotation	2
1.4	1 Differentialgleichung m -ter Ordnung	3
1.5	Grundlagen aus der Analysis	3
2	Das (explizite) Euler-Verfahren	5
3	Einschrittverfahren	8
3.1	Lokaler und globaler Fehler, Rundungsfehler	8
3.1.1	Konsistenz	10
3.1.2	Rundungsfehler	10
3.2	Potenzreihen-Verfahren, Taylorreihen-Verfahren	11
3.3	Runge-Kutta-Verfahren	12
3.3.1	Allgemeines s -stufiges Runge-Kutta-Verfahren	13
3.4	Extrapolations-Verfahren	15
4	Lineare Mehrschritt-Verfahren	15
4.1	Definition	15
4.2	Exkurs zur Lagrange-Interpolationsformel	15
4.3	Exkurs zur Newtonschen Interpolationsformel	16
4.4	Adams-Bashforth-Verfahren	16
4.5	Adams-Moulton-Verfahren (implizit)	17
4.6	Rückwärtsdifferenziationsformeln (BDF)	18
4.7	Die Ordnung eines LMSV	18
4.8	Stabilität	20
4.9	Globaler Fehler, Konvergenz	21
5	Steife Differentialgleichungen, Stabilitätsgebiet	22
5.1	Ordnungsschranke für A-stabile LMSV	26

1 Problemstellung, Notation

1.1 1 Differentialgleichung 1. Ordnung

Definition: Eine *Differentialgleichung 1. Ordnung* ist eine Funktion

$$f : D_f \subset \underbrace{\mathbb{R} \times \mathbb{R}}_{\mathbb{R}^2} \rightarrow \mathbb{R}$$

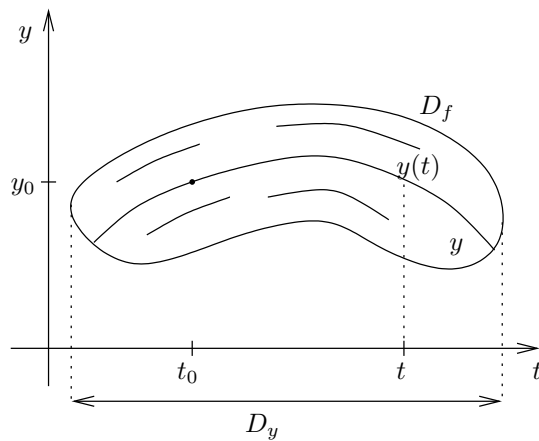
- wobei $D_f \subset \mathbb{R} \times \mathbb{R}$ der Definitionsbereich von f ist - sowie einem *Anfangswert* $(t_0, x_0) \in D_f$.

Eine *Lösung* einer Differentialgleichung 1. Ordnung ist eine Funktion $y : D_y \rightarrow \mathbb{R}$

- wobei $D_y \subset \mathbb{R}$ der Definitionsbereich von y ist - sodass gilt:

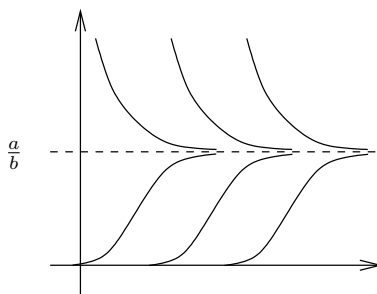
$$\boxed{\frac{dy(t)}{dt} = f(t, y(t))} \quad (1.1a)$$

$$\boxed{y(t_0) = y_0} \quad (1.1b)$$



Beispiel: Biologie: Bevölkerungswachstum

1. $y'(t) = a \cdot y(t)$
 $\Rightarrow y(t) = y_0 \cdot e^{a \cdot (t-t_0)}$
2. $y'(t) = (a - b \cdot y(t)) \cdot y(t)$
 $\Rightarrow y'(t) = 0 \iff y = 0 \text{ oder } y = \frac{a}{b}$



1.2 Systeme von m Differentialgleichungen 1. Ordnung

Definition: Ein System von m Differentialgleichungen 1. Ordnung besteht aus m Funktionen

$$f_k : D_f \subset \underbrace{\mathbb{R} \times \mathbb{R}^m}_{\mathbb{R}^{m+1}} \rightarrow \mathbb{R} \quad k = 1, \dots, m$$

- sowie einem Anfangswert $(t_0, y_{0_1}, y_{0_2}, \dots, y_{0_m})$.

Eine Lösung dieses Systems besteht aus m Funktionen y_k , $k = 1, \dots, m$, sodass gilt:

$$\boxed{\frac{dy_k(t)}{dt} = f_k(t, y_1(t), \dots, y_m(t))} \quad (1.2a)$$

$$\boxed{y_k(t_0) = y_{0_k} \quad k = 1, \dots, m} \quad (1.2b)$$

(1.2a) sind die m Differentialgleichungen und (1.2b) die m Anfangsbedingungen.

Bemerkung:

$$f_k : D_f \subset \underbrace{\mathbb{R} \times \mathbb{R}^m}_{\mathbb{R}^{m+1}} \rightarrow \mathbb{R} \quad k = 1, \dots, m$$

bedeutet, dass

$$\vec{f} : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$$

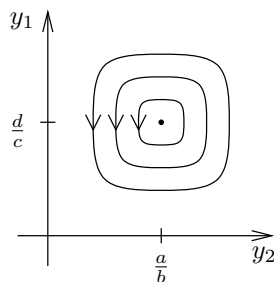
Beispiel: Biologie: Räuber-Beute-Modell

Sei $y_1(t)$ die Grösse einer Kaninchenpopulation und $y_2(t)$ die Grösse einer Fuchspopulation.

$$y_1'(t) = (a - b \cdot y_2(t)) \cdot y_1(t)$$

$$y_2'(t) = (c \cdot y_1(t) - d) \cdot y_2(t)$$

Die stationären Lösungen dieses Problems sind $(y_1, y_2) = (0, 0)$ und $(y_1, y_2) = (\frac{d}{c}, \frac{a}{b})$. Die weiteren Lösungen bewegen sich um die 2. stationäre Lösung herum:



1.3 Vektornotation

Notation: Die Definition eines Systems von m Differentialgleichungen 1. Ordnung lässt sich umformulieren zu

$$\boxed{\frac{dy(t)}{dt} = f(t, y(t))} \quad (1.3a)$$

$$\boxed{y(t_0) = y_0} \quad (1.3b)$$

wobei

$$f := \begin{pmatrix} t \\ f_1 \\ \vdots \\ f_m \end{pmatrix} : D_f \subset \mathbb{R}^{m+1} \rightarrow \mathbb{R}^m : (t, y) \mapsto f(t, y)$$

das Gleichungssystem darstellt, $(t_0, y_0) \in D_f$ der Anfangswert sei und $y : D_y \subset \mathbb{R} \rightarrow \mathbb{R}^m$ die Lösung des Systems ist.

Bemerkung: Diese Schreibweise geht zurück auf Guiseppe Peano (1858 – 1932).

1.4 1 Differentialgleichung m -ter Ordnung

Definition: Sei $m > 1$. Eine Differentialgleichung m -ter Ordnung besteht aus einer Funktion $f : D_f \subset \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}$ – wobei D_f der Definitionsbereich von f ist – sowie m Anfangsbedingungen $y_0^{(0)}, \dots, y_0^{(m-1)}$.

Eine Lösung einer solchen Differentialgleichung ist eine Funktion $y : D_y \subset \mathbb{R} \rightarrow \mathbb{R}$, sodass gilt:

$$\boxed{\frac{d^m y(t)}{dt^m} = f(t, y(t), y^{(1)}(t), \dots, y^{(m-1)}(t))} \quad (1.4a)$$

$$\boxed{y^{(i)}(0) = y_0^{(i+1)} \quad i = 0, \dots, m-1} \quad (1.4b)$$

Bemerkung: Eine Differentialgleichung m -ter Ordnung lässt sich zu einem System von m Differentialgleichungen 1. Ordnung umformulieren:

Sei $y_1(t) := y(t), y_2(t) := y'(t), \dots, y_m(t) := y^{(m-1)}(t)$

$$\Rightarrow \boxed{y_1'(t) = y_2(t), y_2'(t) = y_3(t), \dots, y_m'(t) = f(t, y_1, \dots, y_m)} \quad (1.5a)$$

$$\boxed{y_1(0) = y_0^{(1)}, \dots, y_m(0) = y_0^{(m)}} \quad (1.5b)$$

1.5 Grundlagen aus der Analysis

Satz: Existenz einer Lösung: Peano

Sei $f : D_f \subset \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ stetig.

\Rightarrow Das zu f gehörende System von m Differentialgleichungen 1. Ordnung

1. $\frac{dy(t)}{dt} = f(t, y(t))$
2. $y(t_0) = \vec{y}_0$

besitzt eine Lösung, welche bis an den Rand von D_f fortsetzbar ist.

(Beweis: siehe Analysis)

Satz: Eindeutigkeit der Lösung:

Sei $f : D_f \subset \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ stetig, D_f offen und zusammenhängend und f sei (lokal) Lipschitz-stetig bezüglich y auf allen $K \subset D_f$ kompakt, das heisst

$$\boxed{\forall K \text{ kompakt } \exists L_K < \infty : \|f(t, y) - f(t, \bar{y})\| \leq L_K \|y - \bar{y}\| \quad \forall (t, y), (t, \bar{y}) \in K} \quad (1.6)$$

\Rightarrow Die Lösung des zu f gehörenden Systems von m Differentialgleichungen ist eindeutig.

(Beweis: siehe Analysis)

Definition: Eine Funktion f erfüllt die *einseitige Lipschitz-Bedingung*, wenn es ein Skalarprodukt $\langle \cdot, \cdot \rangle$, die dadurch induzierte Norm $\|\cdot\|$ und ein $l \in \mathbb{R}$ gibt, so dass gilt:

$$\boxed{\langle f(t, y) - f(t, \bar{y}), y - \bar{y} \rangle \leq l \|y - \bar{y}\|^2 \quad \forall (t, y), (t, \bar{y}) \in [t_0, T] \times \mathbb{R}^m} \quad (1.7)$$

Satz: Stetige Abhängigkeit der Lösung vom Anfangswert

Sei $f : [t_0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ stetig, sowie Lipschitz-stetig bezüglich y wie in 1.6 und erfülle die einseitige Lipschitz-Bedingung 1.7. Seien $y, \bar{y} : \mathbb{R} \rightarrow \mathbb{R}^n$ Lösungen von 1.3 zu den Anfangsbedingungen

$$y(t_0) = y_0, \bar{y}(t_0) = \bar{y}_0$$

Dann gilt für alle $t \in [t_0, T]$

$$\|y(t) - \bar{y}(t)\| \leq e^{l(t-t_0)} \cdot \|y_0 - \bar{y}_0\| \quad (1.8)$$

Beweis:

$$\begin{aligned} \frac{d}{dt} \|y(t) - \bar{y}(t)\|^2 &= 2 \langle y'(t) - \bar{y}'(t), y(t) - \bar{y}(t) \rangle \\ &= 2 \langle f(t, y(t)) - f(t, \bar{y}(t)), y(t) - \bar{y}(t) \rangle \\ &\leq 2l \|y(t) - \bar{y}(t)\|^2 \end{aligned}$$

Falls $y(t_0) \neq \bar{y}(t_0) \Rightarrow y(t) \neq \bar{y}(t) \quad \forall t$ (wegen Eindeutigkeit der Lösung)

Also ist $\varphi(t) := \|y(t) - \bar{y}(t)\|^2 \neq 0 \quad \forall t$

$$\Rightarrow \frac{d}{dt} \ln(\varphi(t)) = \frac{\varphi'(t)}{\varphi(t)} \leq 2l$$

Durch Integration folgt:

$$\begin{aligned} \int \frac{\varphi'(t)}{\varphi(t)} dt &\leq \int 2l dt \\ \ln \left(\frac{\varphi(t)}{\varphi(t_0)} \right) &\leq 2l(t - t_0) \\ \|y(t) - \bar{y}(t)\|^2 &\leq e^{2l(t-t_0)} \|y_0 - \bar{y}(0)\|^2 \\ \|y(t) - \bar{y}(t)\| &\leq e^{l(t-t_0)} \|y_0 - \bar{y}(0)\| \end{aligned}$$

$$\Rightarrow \varphi(t) \leq e^{l(t-t_0)} \varphi(t_0)$$

□

Bemerkung:

1. Es ist möglich (und sogar angenehm), dass $l \leq 0$ ist.
2. Falls \exists Lipschitz- $L \Rightarrow \exists$ einseitige Lipschitz $l = L$ o.B.d.A. $l \leq L$
Cauchy-Schwarz-Ungleichung \Rightarrow

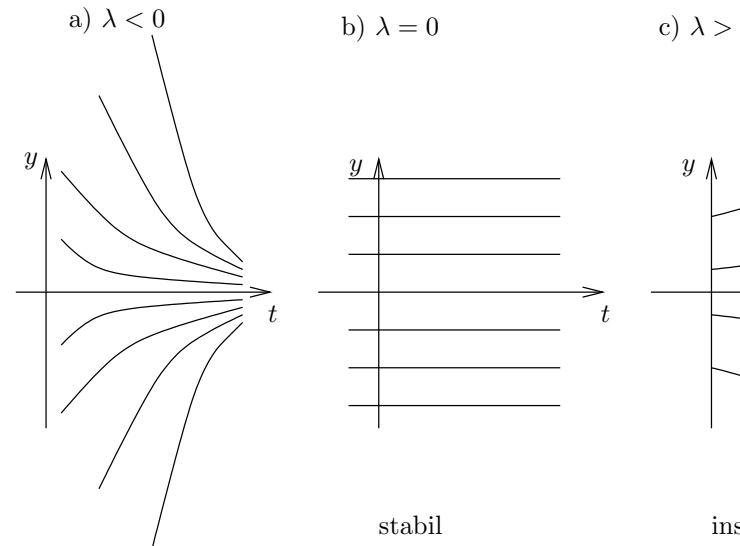
$$\begin{aligned} \|\langle f(t, y(t)) - f(t, \bar{y}(t)), y - \bar{y} \rangle\| &\leq \|f(t, y(t)) - f(t, \bar{y}(t))\| \cdot \|y(t) - \bar{y}(t)\| \\ &\leq L \cdot \|y(t) - \bar{y}(t)\| \end{aligned}$$

Zugleich kann für ein $l \ll L$ gelten:

$$\langle f(t, y) - f(t, \bar{y}), y - \bar{y} \rangle \leq l \|y - \bar{y}\|^2 \quad \forall (t, y), (t, \bar{y}) \in [t_0, T] \times \mathbb{R}^m$$

3. Falls $l < 0$ gilt, laufen die Lösungen zusammen, die Fehler werden kleiner.
(Kontraktion)

Beispiel: $y'(t) = \lambda y(t) \Rightarrow y(t) = y_0 e^{\lambda(t-t_0)}$



Einseitige Lipschitz-Konstante: $l = \lambda$, $L = |\lambda|$ stabil und asymptotisch konvergent

2 Das (explizite) Euler-Verfahren

Bemerkung: Das Verfahren geht zurück auf Leonard Euler (1707-1783).

Bemerkung: Es sei $f : D_f \subset \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ ein System von m Differentialgleichungen mit Anfangswert (t_0, y_0) . Ziel ist es nun, ein Verfahren zu finden, welches diese (unter Umständen nicht oder nur schwer exakt lösbare) Differentialgleichung numerisch löst bzw. approximiert.

Algorithmus: Sei also $D_f = [a, b] \times \mathbb{R}^m$, $f : D_f \rightarrow \mathbb{R}^m$ und $t_0 = a$.

$$\boxed{\begin{cases} \frac{dy(t)}{dt} = f(t, y(t)), & t \in [a, b] \\ y(a) = y_0 \end{cases}} \quad (2.1)$$

Seien $t_i \in [a, b]$ äquidistante Gitterpunkte,

$$t_i := a + i \cdot h \quad i = 0, \dots, N \quad h := \frac{b-a}{N}$$

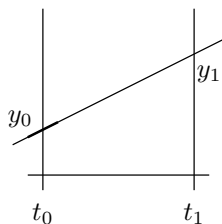
Ziel: Werte $y_i \approx y(t_i)$, ausgehend von $y_0 = y(t_0)$

Idee: Annahme: $y \in C^2([a, b], \mathbb{R}^m)$

$$\boxed{y(t_1) = \underbrace{y_0 + hy'(t_0)}_{=: y_1} + \mathcal{O}(h^2)} \quad (2.2)$$

und analog:

$$\boxed{y_{n+1} := y_n + hy'(t_n) \quad n = 0, \dots, N-1} \quad (2.3)$$



Bemerkung: Wie gross ist der globale Fehler $y_n - y(t_n)$ nach n Schritten? Der Fehler innerhalb eines Schrittes ist $\mathcal{O}(h^2)$; dieser Fehler kann sich allerdings von Schritt zu Schritt fortpflanzen und dadurch auch viel grösser werden als $N \cdot \mathcal{O}(h^2)$! Siehe auch den ersten Schritt vom Beweis von Satz 2.

Bemerkung:

$$\|f(t, y(t)) - f(t, \bar{y}(t))\| \leq L \|y(t) - \bar{y}(t)\| \quad \forall y \in C^1(\mathbb{R}, \mathbb{R}^m) \quad (2.4a)$$

wobei

$$L := \sup_{(t,y) \in D_f} \left\| \frac{\partial f}{\partial y}(t, y) \right\| \leq \sup_{(t,y) \in [a,b] \times \mathbb{R}^m} \left\| \frac{\partial f}{\partial y}(t, y) \right\| < \infty \quad (2.4b)$$

\Rightarrow Die exakte Lösung ist 2mal stetig differenzierbar, denn

$$\begin{aligned} \frac{d}{dt} y(t) &= f(t, y(t)) \\ \frac{d^2}{dt^2} y(t) &= \frac{d}{dt} f(t, y(t)) \\ &= \frac{df}{dt} + \underbrace{\frac{dy(t)}{dt}}_f \cdot \frac{df(t, y)}{dy(t)} \\ \Rightarrow \quad \frac{d^2}{dt^2} y(t) &= \frac{df(t, y(t))}{dt} + f(t, y(t)) \cdot \frac{\partial f}{\partial t}(t, y(t)) \end{aligned} \quad (2.5)$$

Satz: Schranke für den globalen Fehler des Euler-Verfahrens:

Unter obiger Annahme gilt:

$$\|y_n - y(t_n)\| \leq h \frac{e^{L(t_n-a)}}{L} \underbrace{\frac{1}{2} \max_{t \in [a, t_n]} \|y''(t)\|}_{\leq C := \frac{1}{2} \max_{t \in [a, b]} \|y''(t)\|} \quad (2.6)$$

Bemerkung: Also verhält sich der lokale Fehler wie $\mathcal{O}(h^2)$ und der globale Fehler wie $\mathcal{O}(h)$. Ein solches Verfahren heisst *Verfahren 1. Ordnung*, das heisst, es hat *Fehlerordnung* $p = 1$.

Korollar: Konvergenz des Euler-Verfahrens:

Sei $t_n = a + n \cdot h$ fest.

$$\Rightarrow \lim_{\substack{h \rightarrow 0 \\ n \rightarrow \infty}} \|y_n - y(t_n)\| = 0 \quad (2.7)$$

Beweis: Bei festem t_n gilt nach (2.6): $\|y_n - y(t_n)\| \leq h \cdot \text{const} \rightarrow 0$ (unter Vernachlässigung der Rundungsfehler). Die Konvergenz ist sogar gleichmässig (ersetze in der Schranke t_n durch b). \square

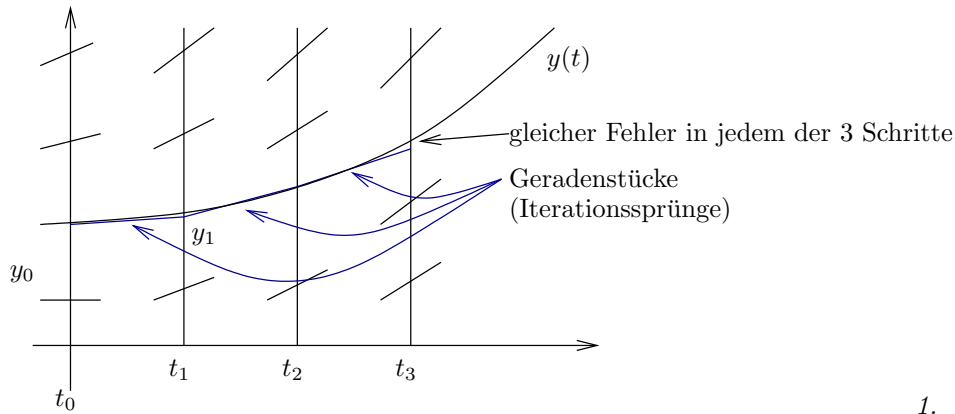
Beweis des Satzes: (in 3 Schritten)

Sei $z(t)$ die Lösung des Anfangswertproblems

$$\begin{cases} \frac{d}{dt} z(t) = f(t, z(t)) \\ z(t_n) = y_n \end{cases} \quad (2.8)$$

Definiere den lokalen Fehler

$$\varphi(t_n, h) := \underbrace{z(t_n + h)}_{t_{n+1}} - y_{n+1} \quad (2.9)$$



1.

Schritt: Taylor mit Integralrestglied für $z \in C^2[a, b]$

$$z(t_n + h) = \underbrace{z(t_n)}_{y_n} + \underbrace{hz'(t_n)}_{hf(t_n, y_n)} + \underbrace{h^2 \int_0^1 (1 - \theta) z''(t_n + \theta h) d\theta}_{\text{lokaler Fehler } \varphi(t_n, h)}$$

Beweis: Sei $s := \theta h \Rightarrow \theta = \frac{s}{h}, \frac{d\theta}{ds} = \frac{1}{h}$. Dann

$$\begin{aligned} & h^2 \int_0^1 (1 - \theta) z''(t_n + \theta h) d\theta \\ &= h^2 \int_0^h \left(1 - \frac{s}{h}\right) z''(t_n + s) \frac{1}{h} ds \\ &= h \cdot z' \Big|_0^h - \int_0^h s z''(t_n + s) ds \\ &= h \cdot z'(t_n + h) - h \cdot z'(t_n) - s \cdot z'(t_n + s) \Big|_0^h + \int_0^h z' ds \\ &= h z'(t_n + h) - h z'(t_n) - h z'(t_n + h) + 0 + z(t_n + h) - z(t_n) \\ &= z(t_n + h) - z(t_n) - h \cdot z'(t_n) \end{aligned}$$

Also gilt

$$\|\varphi(t_n, h)\| \leq h^2 \frac{1}{2} \max_{t \in [t_n, t_{n+1}]} \| \underbrace{z''(t)}_{=f'} \| =: c_n(z) h^2 h^{\frac{1}{2}} \max \left\| \frac{df}{dt} \right\| \quad (2.10)$$

2. Schritt: Fehlerfortpflanzung (pro Schritt)

Betrachte zwei Euler-Schritte vom selben t_n , aber mit verschiedenen Startwerten:

$$v_{n+1} = v_n + hf(t_n, v_n) \text{ und } w_{n+1} = w_n + hf(t_n, w_n)$$

Dann gilt:

$$\begin{aligned} \|v_{n+1} - w_{n+1}\| &\leq \|v_n - w_n\| + h \|f(t_n, v_n) - f(t_n, w_n)\| \\ &\leq \underbrace{(1 + hL)}_{\text{Fehlervverstärkungsfaktor}} \|v_n - w_n\| \quad L = \sup \left\| \frac{df}{dy} \right\| \end{aligned} \quad (2.11)$$

Bemerkung: Eine einseitige Lipschitzbedingung ergäbe eine bessere Schranke.

3. Schritt: Fehlerakkumulation: Lady Windermere's Fächer:

Bemerkung: Sei y_j^k , $k < j$ die Näherung in $y(t_j)$, wenn wir zum Zeitpunkt t_k in $(t_k, y(t_k))$ starten und $j - k$ Schritte machen:

$$y(t_k) = y_k^k \rightsquigarrow y_{k+1}^k \rightsquigarrow y_{k+2}^k \rightsquigarrow \dots \rightsquigarrow y_j^k$$

Insbesondere ist $y_n^0 = y_n$ und $y_k^k = y(t_k)$. Analog zu Schritt 1 gilt:

$$\|y_{k+1}^k - y(t_{k+1})\| \leq c_k(y) h^2 \quad (2.12)$$

wobei c_k lokal abgeschätzt wird durch C :

$$c_k(y) = \frac{1}{2} \max_{t \in [t_k, t_{k+1}]} \|y''(t)\| \leq C = \frac{1}{2} \max_{t \in [a, b]} \|y''(t)\|$$

($y(t)$ ist die gesuchte exakte Lösung des Anfangswertproblems.)

Aus Schritt 2 folgt:

$$\begin{aligned} \|y_n^k - y_n^{k+1}\| &\leq (1 + hL) \|y_{n-1}^k - y_{n-1}^{k+1}\| \leq \dots \leq (1 + hL)^{n-k-1} \|y_{k+1}^k - y_{k+1}^{k+1}\| \\ \Rightarrow \|y_n^k - y_n^{k+1}\| &\leq (1 + hL)^{n-k} \cdot c_k(y) \cdot h^2 \leq (1 + hL)^{n-k} \cdot C \cdot h^2 \end{aligned} \quad (2.13)$$

Somit gilt:

$$\begin{aligned} \|y_n - y(t_n)\| &= \|y_n^0 - y_n^n\| = \|y_n^0 + \sum_{k=1}^{n-1} y_n^k - \sum_{k=1}^{n-1} y_n^k - y_n^n\| \\ &\leq \|y_n^0 - y_n^1\| + \|y_n^1 - y_n^2\| + \dots + \|y_n^{n-1} - y_n^n\| \\ &\leq Ch^2(1 + hL)^{n-1} + Ch^2(1 + hL)^{n-2} + \dots + Ch^2 \\ &\stackrel{Geom.Reihe}{=} Ch^2 \cdot \frac{(1 + hL)^n - 1}{(1 + hL) - 1} \\ &\stackrel{1+hL \leq e^{hL}}{\leq} Ch \frac{e^{nhL} - 1}{L} \\ \Rightarrow \|y_n^0 - y_n^n\| = \|y_n - y(t_n)\| &\leq h \frac{e^{(t_n - a)L} - 1}{L} \cdot C \end{aligned} \quad (2.14)$$

mit

$$\begin{aligned} C &= \frac{1}{2} \sup_{t \in [a, b]} \left\| \frac{d^2}{dt^2} y(t) \right\| = \frac{1}{2} \sup \left\| \frac{d}{dt} f(t, y(t)) \right\| \\ L &= \sup_{D_f} \left\| \frac{d}{dy} f(t, y) \right\| \\ (1 + hL)^n &= \left(1 + \frac{nhL}{n}\right)^n \xrightarrow{n \rightarrow \infty} e^{nhL} \end{aligned}$$

□

3 Einschrittverfahren

3.1 Lokaler und globaler Fehler, Rundungsfehler

Das *Einschrittverfahren* (ESV) ist definiert durch

$$y_{n+1} = y_n + h \cdot \Phi(t_n, y_n, h) \quad (3.1)$$

wobei Φ die Zuwachsfunktion auf $D_f \times [0, H]$ ist.

Notation: Mit y_k^l bezeichnen wir den Wert, der mit dem gegebenen ESV von der exakten Lösung

$$y(l \cdot h) = y(t_l) = y_l^l$$

her in $k - l$ Schritten iteriert wurde, zum Beispiel

$$y_{q+1}^q = y_q^q + h \cdot \Phi(t_q, y_q^q, h)$$

Beispiel: Euler: $\Phi = f$

Definition: Der *lokale Fehler* im Schritt n ist

$$\varphi(t_n, h) = y_{n+1}^{n+1} - y_{n+1}^n$$

Definition: Das ESV hat die *Fehlerordnung* p falls für ein $C > 0$ gilt:

$$\boxed{\| \underbrace{y_{n+1}^{n+1} - y_{n+1}^n}_{=\varphi(t_n, h)} \| \leq Ch^{p+1}} \quad (3.2)$$

Annahme: $\Phi(t, y, h)$ sei Lipschitz-stetig mit der Konstanten $M > 0$ in y , d.h. $\forall t, y, \tilde{y}, h$

$$\boxed{\|\Phi(t, y, h) - \Phi(t, \tilde{y}, h)\| \leq M\|y - \tilde{y}\|} \quad M = \sup \left\| \frac{d}{dy} \Phi(t, y, h) \right\| \quad (3.3)$$

d.h.

$$M = \sup_{D_f \times [0, H]} \left\| \frac{d}{dy} \Phi \right\|$$

falls D_f kompakt und konvex ist.

Satz: Schranke für den globalen Fehler eines ESV der Ordnung p mit Lipschitz-Bedingung

Aus der (3.2) und (3.3) folgt, dass

$$\boxed{\|y_n - y(t_n)\| \leq h^p \frac{e^{M(t_n - a)} - 1}{M} \cdot C} \quad (3.4)$$

mit

$$C = \frac{1}{(p+1)!} \sup_{D_f} \left\| \underbrace{\frac{d^p}{dt^p} f(t, y(t))}_{\frac{d^{p+1}}{dt^{p+1}} y(t)} \right\| \quad (\text{Restglied Taylorreihe})$$

$$M = \sup_{D_f \times [0, H]} \left\| \frac{d}{dy} \Phi(t, y, h) \right\| \quad (\text{Schrankensatz, Analysis II})$$

Beweis: Genau wie beim Euler-Verfahren.

Der lokale Fehler erfüllt (3.2). Also gilt für die Fehlerfortpflanzung:

$$\begin{aligned} \|v_{n+1} - w_{n+1}\| &\leq \|v_n - w_n\| + h \underbrace{\|\Phi(t_n, v_n, h) - \Phi(t_n, w_n, h)\|}_{\leq M\|v_n - w_n\|} \\ &\leq (1 + hM)\|v_n - w_n\| \end{aligned}$$

Also hat man bei der Fehlerakkumulation

$$\begin{aligned} \|y_n^k - y_n^{k-1}\| &\leq (1 + hM)\|y_{n-1}^k - y_{n-1}^{k-1}\| \\ &\leq (1 + hM)^{n-k}\|y_k^k - y_k^{k-1}\| \\ &\leq (1 + hM)^{n-k} \cdot C \cdot h^{p+1} \end{aligned}$$

Genau wie beim Beweis von Satz 2 ist

$$\begin{aligned} \|y_n^0 - y_n^n\| &\leq \sum_t \|y_n^t - y_n^{t+1}\| \\ &\leq \underbrace{\|y_n^0 - y_n^1\|}_{C \cdot h^{p+1}} \cdot \sum_{k=0}^{n-1} (1 + hM)^k \\ &\leq C \cdot h^{p+1} \cdot \frac{(1 + hM)^n - 1}{-1 + (1 + hM)} \\ &\leq \frac{e^{nhM} - 1}{M} \cdot C \cdot h^p \end{aligned}$$

□

3.1.1 Konsistenz

Nun hat der lokale Fehler die Form

$$\begin{aligned}\varphi(t_n, h) &= z(t_n, h) - y_{n+1} \\ &= z(t_n) + h z'(t_n) + \frac{1}{2} h^2 z''(t_n) + \cdots - y_n - h \Phi(t_n, y, h) \\ \Rightarrow \boxed{\varphi(t_n, h) = h(f(t_n, y_n) + \frac{1}{2} h \frac{d}{dt} f(t_n, y(t_n)) + \cdots - \Phi(t_n, y_n, h))} \quad (3.5)\end{aligned}$$

Wir brauchen für $p > 0$:

$$\boxed{f(t_n, y_n) = \Phi(t_n, y_n, 0)} \quad (3.6)$$

Diese Gleichung ist auch bekannt als *Konsistenzbedingung*.

Bemerkung: Begründung: Bei infinitesimalen Schrittweiten

$$\lim_{h \rightarrow 0} \Phi(t, y, h) = \Phi(t, y, 0)$$

entspricht Φ der Ableitung, also f . Der Satz betrifft nur die *Diskretisationsfehler*, nicht die *Rundungsfehler*.

3.1.2 Rundungsfehler

Statt

$$y_{n+1} = y_n + h \Phi(t_n, y_n, h)$$

berechnet man

$$\overline{y_{n+1}} = \overline{y_n} + h \Phi(t_n, \overline{y_n}, h) + \delta_n$$

wobei δ_n der lokale Rundungsfehler ist.

Bemerkung: $\overline{y_k}$ ist nicht der gerundete Wert von y_k , sondern der gerundete Wert + Rundungsfehler vom Verfahren! Annahme: $\|\delta_n\| \leq \delta \forall n$

Also: $\|\overline{y_{n+1}} - y_{n+1}\| \leq (1 + hM)\|\overline{y_n} - y_n\| + \delta$

Lemma: Sei $\{\zeta_n\}_{n \in \mathbb{N}_0}$ eine Folge reeller Zahlen und $A, B \geq 0$, sodass gelte:

$$\boxed{|\zeta_{n+1}| \leq A \cdot |\zeta_n| + B \quad n = 0, 1, \dots, N-1} \quad (3.7)$$

Dann gilt für $n = 1, \dots, N$:

$$\boxed{\zeta_n \leq \begin{cases} A^n |\zeta_0| + \frac{A^n - 1}{A - 1} \cdot B, & A \neq 1 \\ A^n |\zeta_0| + nB, & A = 1 \end{cases}} \quad (3.8)$$

Ist $A = 1 + \varepsilon$ mit $\varepsilon > -1$, so gilt insbesondere

$$\boxed{|\zeta_n| \leq e^{n\varepsilon} |\zeta_0| + \frac{e^{n\varepsilon} - 1}{\varepsilon} \cdot B} \quad (3.9)$$

□

Beweis: siehe Dumode, p. 18

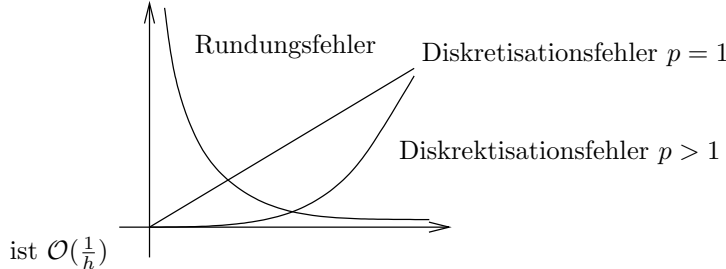
□

Bemerkung: Insbesondere gilt $A = 1 + \varepsilon < e^\varepsilon$.

Hier folgt mit $A := 1 + hM, B := \delta, \zeta_0 := 0$:

$$\|\bar{y}_{n+1} - y_{n+1}\| \leq \frac{\delta}{h} \cdot \frac{e^{nhM} - 1}{M} = \frac{\delta}{h} \cdot \underbrace{\text{const}}_{\text{falls } nh \text{ fest}} \quad (3.10)$$

Das heisst, die Schranke für die in $t_n = a + nh$ akkumulierten Rundungsfehler Fehler in $t_n = an + h$ (fest)



3.2 Potenzreihen-Verfahren, Taylorreihen-Verfahren

Die Formel (3.5) kann man schreiben als

$$\begin{aligned} \varphi(t_n, h) &= z(t_n, h) - y_{n+1} \\ &= h \left(z'(t_n) + \frac{h}{2} z''(t_n) + \dots + \frac{h^{p-1}}{p!} z^{(p)}(t_n) + \text{Restglied } (*) - \Phi(t_n, y_n, h) \right) \end{aligned}$$

Mit $z'(t) = f(t, z(t))$ und $z(t_n) = y_n$ folgt:

$$\varphi(t_n, h) = h \sum_{j=0}^{p-1} \frac{h^j}{(j+1)!} \cdot \frac{d^j}{dt^j} f(t_n, y_n) + \mathcal{O}(h^{p+1}) - h\Phi(t_n, y_n, h)$$

Dies legt nahe zu wählen:

$$\Phi(t_n, y_n, h) := \sum_{j=0}^{p-1} \frac{h^j}{(j+1)!} \frac{d^j}{dt^j} f(t_n, y_n) \quad (3.12)$$

Ein solches Verfahren heisst *Taylorreihen-Verfahren*.

Es wird

$$\varphi(t_n, h) = \mathcal{O}(h^{p+1})$$

das heisst, das Verfahren hat Fehlerordnung p .

Die Ableitungen werden aber kompliziert (doch man könnte Maple oder Mathematica anwenden).

Bemerkung: (*) Restglied:

Das $(p+1)$ -te Restglied ist

$$\int_0^1 \frac{(1-\theta)^p}{p!} z^{(p+1)}(t_n + \theta h) h^p d\theta \quad (3.11)$$

Dies gilt auch, wenn t, h Vektoren sind.

Beispiel: Taylorreihen-Verfahren mit $p = 2$:

$$\begin{aligned} \Phi(t_n, y_n, h) &= f(t_n, y_n) + \frac{h}{2} \left(f_t(t_n, y_n) + \frac{\partial f}{\partial y}(t_n, y_n) \cdot f(t_n, y_n) \right) \\ &= f + \frac{h}{2} (f_t + f_y \cdot f)|_{(t_n, y_n)} \end{aligned} \quad (3.13)$$

Es gibt auch eine Theorie (mit Bäumen), wie man solche Ableitungen kochbuchmässig erzeugt.

3.3 Runge-Kutta-Verfahren

Bemerkung: Wie kommt man zu einfachen Methoden mit $p > 1$?
Es genügt den Fall $m = 1$ (1 Differentialgleichung) zu betrachten.

1. Idee: Euler-Schritt mit h kombinieren mit 2 Schritten mit $\frac{h}{2}$; ergibt

$$\widehat{y}_{n+1} := y_n + \underbrace{\frac{h}{2}f(t_n, y_n)}_Y + \overbrace{\frac{h}{2}f(t_n + \frac{h}{2}, y_n + \frac{h}{2}f(t_n, y_n))}^{h\widehat{\Phi}(t_n, y_n, h)} \quad (3.14)$$

Lokaler Fehler:

$$\widehat{\varphi}(t_n, h) := z(t_n + h) - \widehat{y}_{n+1} = \dots = \frac{h^2}{4}z''(t_n) + \mathcal{O}(h^3) = \mathcal{O}(h^2) \quad (3.15)$$

\Rightarrow Ordnung 1, wie Euler. Für das Euler-Verfahren hatten wir

$$\varphi(t_n, h) := z(t_n + h) - y_{n+1} = \frac{h^2}{2}z''(t_n) + \mathcal{O}(h^3) = \mathcal{O}(h^2) \quad (3.16)$$

Differenz (3.15) – (3.16):

$$\psi(t_n, h) := \widehat{y}_{n+1} - y_{n+1} = \frac{h^2}{4}z''(t_n) + \mathcal{O}(h^3) \quad (3.17)$$

Also gilt:

$$\widehat{y}_{n+1} - y_{n+1} = \widehat{\varphi}(t_n, h) + \mathcal{O}(h^3)$$

Das heisst $\widehat{y}_{n+1} - y_{n+1}$ ist gleich dem lokalen Fehler des Doppelschritt-Euler-Verfahrens $+\mathcal{O}(h^3)$

$$\Rightarrow \widehat{y}_{n+1} - y_{n+1} = \frac{1}{2}\varphi(t_n, h) + \mathcal{O}(h^3) \quad (3.18)$$

Also ist $\widehat{y}_{n+1} - y_{n+1}$ gleich $\frac{1}{2}$ mal der lokale Fehler des Euler-Verfahrens $+\mathcal{O}(h^3)$
Dies erlaubt es den lokalen Fehler zu schätzen. \leadsto Schrittweitensteuerung (siehe Jeltsch, p. 163, oder Lubich, § 6.5) Wir können aber (3.15) und (3.16) so konstruieren, dass der Term $h^2z''(t_n)$ wegfällt:

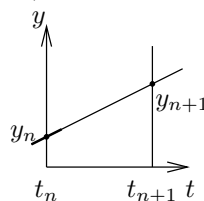
$$z(t_n + h) - \underbrace{(2\widehat{y}_{n+1} - y_{n+1})}_{:=Y_{n+1}; \text{ Euler}} = 2\widehat{\varphi}(t_n, h) - \varphi(t_n, h) = \mathcal{O}(h^3) \quad (3.19)$$

definiert ein neues Verfahren mit Ordnung 2.

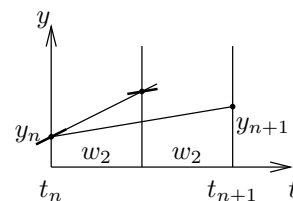
Einsetzen ergibt:

$$y_{n+1} := 2\widehat{y}_{n+1} - y_{n+1} = y_n + \underbrace{hf(t_n + \frac{h}{2}, y_n + \frac{h}{2}f(t_n, y_n))}_{h\Phi(t_n, y_n, h)} \quad (3.20)$$

modifiziertes Euler-Verfahren ($p = 2$)



Euler-Verfahren



modif. Euler-Verfahren

Erfordert 2 Auswertungen von f :

Verfahren von Runge-Kutta

2. Idee: Ansatz der (3.14) verallgemeinert (der Einfachheit halber für $m = 1$):

$$\boxed{\Phi(t, y, h) := bf(t, y) + cf(t + \beta h, y + \gamma hf(t, y))} \quad (3.21)$$

wobei (t, y) der Anfangspunkt und $t + \beta h, y + \gamma hf(t, y)$ ein Zwischenpunkt ($0 \leq \beta \leq 1$) nach "Eulerartigem" Schritt sind (bei Euler ist $\beta = \gamma$).

Gesucht sind b, c, β, γ , sodass die Ordnung p maximal wird.

Taylorentwicklung:

$$f(t + \beta h, y + \gamma hf(t, y)) = f(t, y) + \beta h f_t(t, y) + \gamma h f(t, y) f_y(t, y) + \mathcal{O}(h^2)$$

Also:

$$\begin{aligned} \Phi(t, y, h) &= (b + c)f(t, y) + ch(\beta f_t(t, y) + \gamma f(t, y) f_y(t, y)) + \mathcal{O}(h^2) \\ &= (b + c)f + ch(\beta f_t + \gamma f f_y)|_{(t, y)} \end{aligned}$$

\Rightarrow Übereinstimmung von Φ mit der Taylormethode, falls

$$b + c = 1, \quad c\beta = \frac{1}{2}, \quad c\gamma = \frac{1}{2}$$

\Rightarrow Lösungen für $c \neq 0$ beliebig:

$$b = 1 - c, \quad \beta = \gamma = \frac{1}{2c}$$

Also: Schar von Verfahren mit $p = 2$:

$$\boxed{\Phi(t, y, h) := (1 - c)f(t, y) + cf(t + \frac{h}{2c}, y + \frac{h}{2c}f(t, y))} \quad (3.22)$$

Spezialfälle:

1. $c = 1$: modifiziertes (verbessertes) Euler-Verfahren, Verfahren von Runge-Kutta
2. $c = \frac{1}{2}$: vereinfachtes Runge-Kutta-Verfahren

Auch ok, wenn $m > 1$ (d.h. mit mehreren Differentialgleichungen); Verfahren = Formel von Heun

3.3.1 Allgemeines s -stufiges Runge-Kutta-Verfahren

Explizite Runge-Kutta-Verfahren:

$$\boxed{\mathbf{y}_i = \mathbf{y}_n + h \sum_{j=1}^{i-1} a_{ij} f(t_n + c_j h, \mathbf{y}_j), \quad i = 1, \dots, s} \quad (3.23a)$$

$$\boxed{y_{n+1} = y_n + h \sum_{j=1}^s b_j f(t_n + c_j h, \mathbf{y}_j)} \quad (3.23b)$$

Schreiben die Koeffizienten ins *Butcher-Tableau*:

$c_1 = 0$	0	0
c_2	a_{21}	0	\ddots		\vdots
\vdots	\vdots		\ddots		\vdots
c_i	a_{i1}	a_{i2}	...	$a_{i,i-1}$	\ddots
\vdots	\vdots	\vdots		\ddots	\vdots
c_s	a_{s1}	a_{s2}	...	$a_{s,i-1}$...
	b_1	b_2	...	b_{i-1}	...
				b_{s-1}	b_s

Bedingungen:

$$\boxed{c_i = \sum_{j=1}^{i-1} a_{ij} = \sum_{j=1}^s a_{ij} \quad (\text{ohne Beweis})} \quad (3.24a)$$

$$\boxed{\sum_{i=1}^s b_i = 1 \quad (\text{Formel exakt für } f(t) = 1)} \quad (3.24b)$$

Bemerkung: Bei expliziten Verfahren ist die Diagonale $(a_{i,i})_{i=1,\dots,s}$ 0.

Beispiel: Butcher-Tableaus für:

1. Explizites Euler-Verfahren, $s = 1$:

$$\begin{array}{c|c} 0 & 0 \\ \hline & 1 \end{array} \mathbf{y}_1 = y_n$$

2. modifiziertes Euler / Runge, $s = 2$:

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \hline & 0 & 1 \end{array} \mathbf{y}_1 = y_n, \mathbf{y}_2 = y_n + \frac{h}{2} f(t_n, y_n)$$

3. Heun, $s = 2$:

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

4. Verfahren (3.22), $s = 2$:

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \frac{1}{2c} & \frac{1}{2c} & 0 \\ \hline & 1-c & c \end{array}$$

5. klassisches Runge-Kutta: $s = 4$:

$$\begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ \hline & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{array}$$

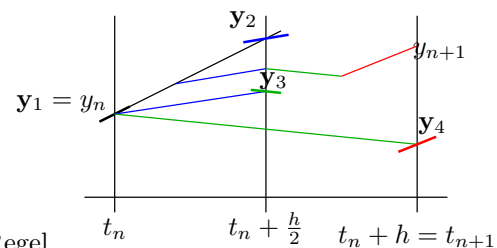
$$\mathbf{y}_1 = y_n$$

$$\mathbf{y}_2 = y_n + \frac{h}{2} f(t_n, \mathbf{y}_1)$$

$$\mathbf{y}_3 = y_n + \frac{h}{2} f(t_n + \frac{h}{2}, \mathbf{y}_2)$$

$$\mathbf{y}_4 = y_n + h f(t_n + \frac{h}{2}, \mathbf{y}_3)$$

$$y_{n+1} = y_n + h \left(\frac{1}{6} f(t_n, \mathbf{y}_1) + \frac{1}{3} f(t_n + \frac{h}{2}, \mathbf{y}_2) + \frac{1}{3} f(t_n + \frac{h}{2}, \mathbf{y}_3) + \frac{1}{6} f(t_n + h, \mathbf{y}_4) \right) \quad (3.25)$$



Ordnung 4, im Falle $f(t, y) = f(t)$: Simpson-Regel.

Erreichbare Fehlerordnung bei expliziten Runge-Kutta-Verfahren:

p	1	2	3	4	5	6	7	9	10
s	1	2	3	4	6	7	9	12	13

Implizite Runge-Kutta-Verfahren: Im Prinzip könnten alle Elemente der $s \times s$ -Matrix $(a_{ij}) \neq 0$ sein, aber in der Regel wählt man nur zusätzliche Elemente $\neq 0$ auf der Diagonalen, das heisst $a_{jj} \neq 0$ ($j = 1, \dots, s$).

Implizites Euler-Verfahren: $s = 1$:
$$\begin{array}{c|c} 1 & 1 \\ \hline & 1 \end{array}$$

$$\mathbf{y}_1 = y_n + h f(t_n + h, \mathbf{y}_1)$$

$$y_{n+1} = y_n + hf(t_n + h, \mathbf{y}_1)$$

$\Rightarrow y_{n+1} = y_1$, also:

$$\boxed{y_{n+1} = y_n + hf(t_{n+1}, y_{n+1})} \quad (3.26)$$

Für die Schrittweitensteuerung nimmt man am besten zwei Runge-Kutta-Verfahren, deren Ordnung um 1 verschieden ist, die aber dieselben c_i und a_{ij} verwenden, ausser dass das y_{n+1} der einen Methode noch als y_s der zweiten Methode verwendet wird (\leadsto eingebettete Runge-Kutta-Formeln).

Ein Beispiel mit $s = 6$ bzw. $s = 7$, die Formeln von Dormand und Prince, findet man auf einer Beilage.

3.4 Extrapolations-Verfahren

siehe Bemerkung in *Jeltsch, p.133*, oder *Stoer, Abschnitte 7.2, 12 und 13*, ausserdem *Bulirsch, NM II, 1973. Analogie zur Romberg-Quadratur*.

4 Lineare Mehrschritt-Verfahren

4.1 Definition

Definition: *Lineare Mehrschritt-Verfahren (LMSV)* sind von der Form

$$\boxed{\sum_{j=0}^k \alpha_j y_{n-j} = h \sum_{j=0}^k \beta_j \underbrace{f(t_{n-j}, y_{n-j})}_{=: f_{n-j}} \quad n = k, k+1, \dots} \quad (4.1)$$

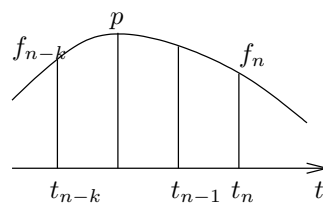
mit festen Koeffizienten $\alpha_0, \dots, \alpha_k, \beta_0, \dots, \beta_k$, wobei $\alpha_0 \neq 0$ und ohne Beschränkung der Allgemeinheit ist $\alpha_k \neq 0$ oder $\beta_k \neq 0$. Weiterhin ist $t_j = t_0 + jh$. Man verwendet "alte" y -Werte und nimmt an, dass y_0, \dots, y_{k-1} durch ein anderes Verfahren berechnet worden sind.

Definition: Man unterscheidet 2 Fälle:

1. $\beta_0 = 0 \Rightarrow y_n$ kommt auf der rechten Seite der Gleichung nicht vor: Explizites LMSV
2. $\beta_0 \neq 0 \Rightarrow y_n$ kommt auf der rechten Seite der Gleichung vor: Implizites LMSV. (In jedem Schritt muss ein Gleichungssystem gelöst werden, meist nicht-linear.)

Bemerkung: Im Allgemeinen sind implizite Verfahren stabiler.

4.2 Exkurs zur Lagrange-Interpolationsformel



Wir wollen (t_{n-j}, f_{n-j}) , $j = 0, \dots, k$ durch ein Polynom p vom Grad k interpolieren, d.h. $p \in \mathcal{P}_k$

$$\boxed{p(t) = \sum_{j=0}^k l_j(t) f_{n-j} \quad l_j \in \mathcal{P}_k \quad l_j(t_{n-m}) = \begin{cases} 1, & j = m \\ 0, & j \neq m \end{cases}} \quad (4.2)$$

$$\boxed{l_j(t) = \frac{\prod_{\substack{l=0 \\ l \neq j}}^k (t - t_{n-l})}{\prod_{\substack{l=0 \\ l \neq j}}^k (t_{n-j} - t_{n-l})}} \quad (4.3)$$

Diese Formel ist allerdings numerisch nicht sehr gut. Wenn man stattdessen interpoliert von (t_{n-j}, f_{n-j}) , $j = 1, \dots, k$:

$$\sum_{j=0}^k \sim \sum_{j=1}^k \quad \prod_{\substack{l=0 \\ l \neq j}}^k \sim \prod_{\substack{l=1 \\ l \neq j}}^k$$

Bei äquidistanten Stützstellen wird der Nenner h^{k-1} bzw. h^k mal eine ganze Zahl.

4.3 Exkurs zur Newtonschen Interpolationsformel

(äquidistant, mit Rückwärtsdifferenzen)

Wir wollen (t_{n-j}, f_{n-j}) , $j = 0, \dots, k$ durch ein Polynom $p \in \mathcal{P}_k$ interpolieren.

Rückwärtsdifferenzen:

$$\begin{aligned} \nabla f_n &= f_n - f_{n-1} \\ \nabla^m f_n &= \nabla(\nabla^{m-1} f_n) = \sum_{j=0}^m \binom{m}{j} (-1)^j f_{n-j} \end{aligned} \quad (4.4)$$

Satz: Das Polynom

$$\begin{aligned} p(t) &:= \sum_{m=0}^k (-1)^m \binom{-\tau}{m} \nabla^m f_n, \\ \text{wo } \tau &:= \frac{t-t_n}{h} \in \mathbb{R} \quad \binom{-\tau}{m} := \frac{-\tau(-\tau-1)\dots(-\tau-m+1)}{m!} \end{aligned} \quad (4.5)$$

interpoliert die Daten (t_{n-j}, f_{n-j}) , $j = 0, \dots, k$

Bemerkung: $p(t)$ hat Grad $\leq k$ bei $k+1$ interpolierten Punkten.

Bemerkung: Dieses Polynom p ist das gleiche wie bei Lagrange, da Polynome vom Grad k durch $k+1$ Stützstellen eindeutig bestimmt sind.

Bemerkung: Anwendungen: Adams-Bashforth-Formeln, Adams-Moulton-Formeln, BDF-Formeln, ...

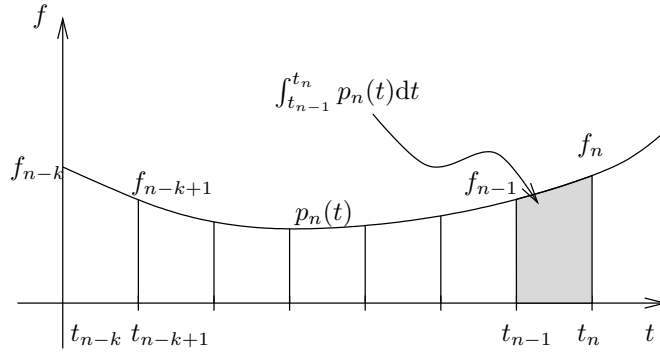
4.4 Adams-Bashforth-Verfahren

Wir legen Interpolationspolynome p_n vom Grad $k-1$ durch die k Punkte

$$(t_{n-j}, \underbrace{f(t_{n-j}, y_{n-j})}_{=: f_{n-j}}), j = 1, \dots, k \quad (4.6)$$

und integrieren dieses exakt zwischen t_{n-1} und t_n .

$$\begin{aligned} y_n &:= y_{n-1} + \int_{t_{n-1}}^{t_n} p_n(t) dt \\ &\approx y(t_{n-1}) + \underbrace{\int_{t_{n-1}}^{t_n} \underbrace{f(t, y(t))}_{y'(t)} dt}_{\underbrace{y(t_n) - y(t_{n-1})}_{y(t_n)}} \end{aligned} \quad (4.7)$$



Eigentlich ist das eine Extrapolation, denn p_n wird ausserhalb der Stützstellen integriert. Nach Lagrange:

$$\begin{aligned} \int_{t_{n-1}}^{t_n} p_n(t) dt &= \int_{t_{n-1}}^{t_n} \sum_{j=1}^k l_j(t) f_{n-j} dt \\ &= \sum_{j=1}^k f_{n-j} \underbrace{\int_{t_{n-1}}^{t_n} l_j(t) dt}_{=: h\beta_j} \end{aligned}$$

d.h. in 4.1 wird

$$\begin{aligned} \alpha_1 &= 1 & \alpha_1 &= -1 & \alpha_2 &= \dots = \alpha_k = 0 \\ \beta_0 &= 0 & \beta_j &= \frac{1}{h} \int_{t_{n-1}}^{t_n} \frac{\prod_{l=1, l \neq j}^k (t - t_{n-l})}{\prod_{l=1, l \neq j}^k (t_{n-j} - t_{n-l})} dt & j &= 1, \dots, k \end{aligned} \quad (4.8)$$

Besser verwendet man die Newtonsche Formel. Man kann zeigen, dass

$$y_n = y_{n-1} + h \sum_{m=0}^{k-1} \gamma_m \nabla^m f_{n-1} \quad \gamma_m := (-1)^m \int_0^1 \binom{-\tau}{m} d\tau \quad (4.9)$$

mit $\gamma_0 = 1$, $\gamma_1 = \frac{1}{2}$, $\gamma_2 = \frac{5}{12}$, $\gamma_3 = \frac{3}{8}$, $\gamma_4 = \frac{251}{720}$, $\gamma_5 = \frac{95}{288}$, unabhängig von k !
Zusammenhang:

$$\beta_{j+1} = (-1)^j \sum_{l=j}^{k-1} \binom{l}{j} \gamma_l, \quad j = 0, \dots, k-1 \quad (4.10)$$

Bemerkung: Dieses Verfahren hat die Fehlerordnung k .

4.5 Adams-Moulton-Verfahren (implizit)

Wir wählen p_n von Grade k statt $k-1$ und interpolieren zusätzlich in

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

Da y_n unbekannt ist, ist das Verfahren implizit, d.h. es muss wieder in jedem Schritt ein im Allgemeinen nicht lineares Gleichungssystem gelöst werden. Oft wird eine erste Näherung von y_n mit Adams-Bashforth bestimmt ("Prädiktor") und dann dieser im entsprechenden Adams-Moulton-Schritt in $f(t_n, y_n)$ korrigiert ("Korrektor"). Nur ein oder wenige Schritte einer Fixpunkt-Iteration werden durchgeführt. Mit Newton-Interpolation:

$$y_n - y_{n-1} = h \sum_{m=0}^k \gamma_m^* \nabla^m f_n \quad \gamma_m^* := (-1)^m \int_0^1 \binom{1-\tau}{m} d\tau \quad (4.11)$$

Fehlerordnung: $k+1$

$k=0$: $p_n(t) \equiv f_n = f(t_n, y_n) \rightsquigarrow$ implizites Euler-Verfahren.

$k=1$: $p_n(t)$ ist eine lineare Funktion durch (t_{n-1}, f_{n-1}) und (t_n, f_n)

$$y_n = y_{n-1} + \frac{h}{2} (f(t_{n-1}, y_{n-1}) + f(t_n, y_n)) \quad (4.12)$$

Trapezregel

Bemerkung: Ist f linear in y , so ist in jedem Schritt nur ein lineares Gleichungssystem im m Unbekannten zu lösen ($y_n \in \mathbb{R}^m$), wobei

$$\gamma_0^* = 1, \quad \gamma_1^* = -\frac{1}{2}, \quad \gamma_2^* = -\frac{1}{12}, \quad \gamma_3^* = -\frac{1}{24}, \quad \gamma_4^* = -\frac{19}{720}, \quad \gamma_5^* = -\frac{3}{160}$$

4.6 Rückwärtsdifferenziationsformeln (BDF)

Hier interpoliert man *nicht* $(t_{n-j}, f_{n-j}) = (t_{n-j}, y'_{n-j})$, sondern direkt die approximierte Lösung

$$(t_{n-j}, y_{n-j}), \quad j = 0, \dots, k$$

Allerdings kennt man y_n gar nicht!

Nach Newton:

$$p(t) = \sum_{m=0}^k (-1)^m \binom{-\tau}{m} \nabla^m y_n, \quad \tau = \frac{t - t_n}{h}$$

$$\boxed{p'(t) = \sum_{m=0}^k (-1)^m \underbrace{\frac{d}{dt} \binom{-\tau}{m}}_{\frac{1}{h} \frac{d}{d\tau} \binom{-\tau}{m}} \nabla^m y_n} \quad (4.13)$$

Zudem fordern wir für $p'(t_n)$ die *Kollokationsbedingung*:

$$\boxed{p'(t_n) = f_n := f(t_n, y_n)} \quad (4.14)$$

Die letzten beiden Gleichungen definieren das *BDF-Verfahren*.

Beispiel: $k = 1$

$$p(t) = y_n + \overbrace{\frac{y_n - y_{n-1}}{h}}^{\nabla y_n / h} (t - t_n)$$

$$p'(t) = \frac{y_n - y_{n-1}}{h} \stackrel{\text{Koll.-Bed.}}{=} f(t_n, y_n)$$

$\Rightarrow y_n = y_{n-1} + hf(t_n, y_n)$, implizites Euler-Verfahren

Beispiel: $k = 2$

$$p(t) = y_n + \frac{1}{h}(t - t_n)\nabla y_n + \frac{1}{2h^2}(t - t_n)(t - t_{n-1})\nabla^2 y_n$$

$$p'(t) = \frac{1}{h}\nabla y_n + \frac{1}{2h^2}(t - t_{n-1} + t - t_n)\nabla^2 y_n$$

$$\Rightarrow p'(t_n) = \frac{1}{h}\nabla y_n + \frac{1}{2h}\nabla^2 y_n = \frac{1}{h}\left(\frac{3}{2}y_n - 2y_{n-1} + \frac{1}{2}y_{n-2}\right)$$

$$\Rightarrow \boxed{y_n - \frac{4}{3}y_{n-1} + \frac{1}{3}y_{n-2} = \frac{2}{3}hf(t_n, y_n)} \quad (4.15)$$

Bemerkung: Für $k \geq 7$ sind die BDF nicht stabil.

4.7 Die Ordnung eines LMSV

Definition: Das LMSV (4.1) hat die Ordnung p , wenn für jedes Anfangswertproblem $y' = f(t, y)y(t_0) = y_0$ mit $f \in C^{p+1}$ gilt, dass bei exakten Startwerten $y_j = y(t_j), j = 0, \dots, k-1$ gilt:

$$\boxed{y_k^{k-1} - y_k^k = \mathcal{O}(h^{p+1})} \quad (4.16)$$

wobei $y_k^{k-1} - y_k^k = y_k - y(t_k)$ der lokale Fehler im ersten Schritt ist.

Hilfssatz: Ein LMSV hat die Ordnung p genau dann, wenn mit

$$E(y, t, h) := \sum_{j=0}^k (\alpha_{k-j} y(t+jh) - h\beta_{k-j} y'(t+jh)) \quad (4.17)$$

für alle $y \in C^{p+1}$ und für alle $t \in \mathbb{R}$

$$E(y, t, h) = \mathcal{O}(h^{p+1}) \quad (4.18)$$

Definition:

$$\frac{E(y, t, h)}{h \sum_{j=0}^k \beta_j}$$

heisst "truncation error".

Beweis: Sei y Lösung. Wegen $y_j = y(t_j), j = 0, \dots, k-1$ gilt

$$\begin{aligned} E(t, y, h) &= \sum_{j=0}^k (\alpha_{k-j} (y(t+jh)) - h\beta_{k-j} \cdot y'(t+jh)) \\ &= \sum_{j=0}^{k-1} (\alpha_{k-j} y_j - h\beta_{k-j} f_j) + \alpha_0 y(t+kh) - h\beta_0 - \\ &\quad - y'(t+kh) + \underbrace{(\alpha_0 y_k - h\beta_0 f_k) - (\alpha_0 y_k - h\beta_0 f_k)}_{=0} \\ &= \sum_{j=0}^k \underbrace{(\alpha_j y_{k-j} - h\beta_j f_{k-j})}_{=0 \text{ nach Def. LMSV}} + (\alpha_0 y(t+kh) - \\ &\quad - h\beta_0 y'(t+kh)) - (\alpha_0 y_k - h\beta_0 f_k) \\ &= \alpha_0 (y(t+kh) - y_k) - h\beta_0 (f(t_k, y(t_k)) - f_k) \\ \Rightarrow \|E(t, y, h)\| &= \|\alpha_0 (y(t+kh) - y_k) - \underbrace{h\beta_0 (f(t_k, y(t_k)) - f_k)}_{\|\cdot\| \leq h \cdot \beta_0 \cdot L \cdot \|y(t_k) - y_k\|}\| \\ &\leq (|\alpha_0| + h|\beta_0|L) \underbrace{\|y(t_k) - y_k\|}_{\in \mathcal{O}(h^{p+1})} \end{aligned}$$

Falls die Ordnung p ist, so ist $\|y(t_k) - y_k\| = \mathcal{O}(h^{p+1})$ und $\|f(t_k, y(t_k)) - f(t_k, y_k)\| \leq L\|y(t_k) - y_k\| \leq \mathcal{O}(h^{p+1}) \Rightarrow (4.18)$

Falls (4.18) $\Rightarrow \mathcal{O}(h^{p+1}) = E(y, t, h) \geq (\alpha_0 - h|\beta_0|L)\|y(t_k) - y_k\| \Rightarrow \|y(t_k) - y_k\| = \mathcal{O}(h^{p+1})$ wie in (4.16) gefordert. \square

Satz: LMSV haben mindestens Ordnung p genau dann, wenn

$$\sum_{j=0}^k \alpha_j = 0 \quad \sum_{j=0}^k \alpha_{k-j} j^l = l \sum_{j=0}^k \beta_{k-j} j^{l-1} \quad (l = 1, \dots, p) \quad (4.19)$$

Bemerkung: Wenn die 2. Beziehung falsch ist für $p+1$, dann ist die Ordnung genau p .

Bemerkung: (4.19) \iff Differentialgleichungen bzw. Integrale $y' = 0$ und $y' = lt^{l-1}$ ($l = 1, \dots, p$) werden exakt gelöst mit $h = 1$ ($\Rightarrow y(t) = t^l, l = 0, \dots, p$).

Beweis:

$$\begin{aligned}
& E(y, t, h) \\
&= \sum_{j=0}^k (\alpha_{k-j} y(t + jh) - h \beta_{k-j} y'(t + jh)) \\
&= \sum_{j=0}^k \left(\alpha_{k-j} \sum_{l=0}^p y^{(l)}(t) \frac{(jh)^l}{l!} - h \beta_{k-j} \sum_{s=0}^{p-1} y^{(s+1)}(t) \frac{(jh)^s}{s!} \right) + \\
&\quad + \mathcal{O}(h^{p+1}) \\
&\stackrel{s+1=l}{=} \left(\sum_{j=0}^k \alpha_j y(t) + \sum_{l=1}^p \frac{h^l}{l!} y^{(l)}(t) \left(\sum_{j=0}^k \alpha_{k-j} j^l - l \sum_{j=0}^k \beta_{k-j} j^{l-1} \right) \right) + \\
&\quad + \mathcal{O}(h^{p+1}) \\
&= \mathcal{O}(h^{p+1}) \iff (4.19)
\end{aligned}$$

□

Bemerkung: Man könnte nun folgendes zeigen:

k -Schritt Adams-Bashforth: Ordnung k
 k -Schritt Adams-Moulton : Ordnung $k + 1$
 k -Schritt BDF : Ordnung k

Definition: Ein LMSV heisst *konsistent*, wenn es Ordnung $p \geq 1$ hat.

Korollar: LMSV konsistent \iff

$$\boxed{\sum_j \alpha_j = 0 \quad \sum_j j \alpha_{k-j} = \sum_j \beta_{k-j}} \quad (4.20)$$

Definition:

$$\boxed{\rho(\zeta) := \sum_{j=0}^k \alpha_j \zeta^{k-j} \quad \sigma(\zeta) := \sum_{j=0}^k \beta_j \zeta^{k-j}} \quad (4.21)$$

Man nennt ρ das 1. charakteristische Polynom und σ das 2. charakteristische Polynom.

Korollar: LMSV konsistent \iff

$$\boxed{\rho(1) = 0 \quad \rho'(1) = \sigma(1)} \quad (4.22)$$

4.8 Stabilität

Definition: LMSV heisst *0-stabil*, wenn jede Lösung der Rekursion

$$\boxed{\alpha_0 y_n + \alpha_1 y_{n-1} + \dots + \alpha_k y_{n-k} = 0, \quad n = k, k+1, \dots} \quad (4.23)$$

beschränkt bleibt für beliebige Anfangswerte y_0, \dots, y_{k-1} und $n \rightarrow \infty$.

Satz: LMSV ist genau dann 0-stabil, wenn für die Nullstelle ζ_i von ρ gilt:

- i) $|\zeta_i| \leq 1, \quad i = 1, \dots, k$
- ii) $|\zeta_i| < 1$, falls ζ_i eine mehrfache Nullstelle ist.

Beweis: Idee:

$y_n = \zeta_i^n$ ist Lösung von (4.23), da

$$\alpha_0 \zeta_i^n + \alpha_1 \zeta_i^{n-1} + \dots + \alpha_k \zeta_i^{n-k} = \zeta_i^{n-k} \rho(\zeta_i) = 0$$

Falls ζ_i eine m -fache Nullstelle ist, ist für $l < m$ auch

$$y_{n,l} := \frac{d^l}{d\zeta^l} \zeta^n \Big|_{\zeta=\zeta_i} = n \cdot (n-1) \dots (n-l+1) \zeta_i^{n-l}$$

eine Lösung, denn

$$\begin{aligned} \sum_{j=0}^k \alpha_j (n-j) \dots (n-j-l+1) \zeta_i^{n-j-l} &= \sum_{j=0}^k \alpha_j \frac{d^l}{d\zeta^l} \zeta^{n-j} \Big|_{\zeta=\zeta_i} \\ &= \frac{d^l}{d\zeta^l} (\zeta^{n-k} \rho(\zeta)) \Big|_{\zeta=\zeta_i} = 0 \end{aligned}$$

Diese k Lösungen zusammen bilden eine Basis des linearen Lösungsraums. Die allgemeine Lösung bleibt beschränkt genau dann, wenn i) und ii) gelten. \square

Beispiel: Adams-Verfahren:

$$\rho(\zeta) = \zeta^k - \zeta^{k-1} = \zeta^{k-1}(\zeta - 1)$$

\Rightarrow 0-stabil.

Man kann zeigen, dass BDF stabil ist für $k \leq 6$ und instabil für $k \geq 7$. Welche Ordnung lässt sich erreichen?

Satz: Dahlquist, 1956

(ohne Beweis)

Für ein 0-stabiles LMSV der Ordnung p und Schrittzahl k gilt:

$p \leq k+1$, falls k ungerade

$p \leq k+2$, falls k gerade

Falls $p = k+2$, so haben alle Nullstellen ζ_i Betrag 1.

4.9 Globaler Fehler, Konvergenz

Satz: Schranke für globalen Fehler

Voraussetzung: AWP mit f genügend diff'bar und Startwerte $y_j, j = 0, \dots, k-1$, sodass

$$\|y_j - y(t_j)\| \leq C \cdot h^p, \quad (j = 0, \dots, k-1) \quad (4.24)$$

LMSV 0-stabil und von Ordnung $p \geq 1$ (konsistent). Behauptung: Für den globalen Fehler gilt die Abschätzung

$$\|y_n - y(t_n)\| \leq M \cdot h^p \quad (4.25)$$

falls $t_n \leq T$, wobei M unabhängig von h und $n, nh \leq T - t_0$

Beweis: Idee:

Transformation auf 1-Schritt-Verfahren (analog Reduktion einer linearen Differentialgleichung höherer Ordnung auf ein System 1. Ordnung), dann Beweis ähnlich wie für Euler- und Runge-Kutta-Verfahren. \square

Definition: LMSV ist *konvergent*, wenn für alle AWP, welche die Voraussetzungen vom Satz über die Eindeutigkeit der Lösung (siehe (1.6) mit $D_f = [t_0, T] \times \mathbb{R}^n$ erfüllen und für alle Anfangswerte y_0, \dots, y_{k-1}

$$\lim_{h \rightarrow 0} y_j = y_0 \quad (j = 1, \dots, k-1) \quad (4.26)$$

gilt:

$$\lim_{\substack{h \rightarrow 0 \\ nh = t - t_0}} y_n = y(t_n) \quad \forall t \in [t_0, T] \quad (4.27)$$

Klar: (4.25), $p \geq 1 \Rightarrow$ (4.27)

(4.24), $p \geq 1 \Rightarrow$ (4.26)

Also: (4.24) + 0-stabil + konsistent \Rightarrow (4.27)

Wäre (4.26) \iff (4.24) mit $p = 1$, dann könnten wir schliessen: 0-stabil + konsistent \Rightarrow konvergent.

Dies gilt effektiv, ebenso die Umkehrung:

Satz: Dahlquist:

Für LMSV gilt:

$$0\text{-stabil} + \text{konsistent} \iff \text{konvergent}$$

Der vollständige Beweis ist sehr aufwändig.

Bemerkung: Diese Aussage gilt (in ähnlicher Weise) auch für andere Verfahren.

5 Steife Differentialgleichungen, Stabilitätsgebiet

Bemerkung: In vielen Anwendungen (z.B. in der Chemie) kommen Systeme von Differentialgleichungen vor, wo die Lösung eine Kombination aus sehr langsam und sehr schnell abfallenden Komponenten ist. Eine solche Differentialgleichung nennt man *steif* [engl.: stiff].

Vereinfachte Situation: lineares System mit konstanten Koeffizienten

$$y' = Ay \quad A \text{ zeitunabhängig}$$

$$y(0) = y_0$$

Annahme: A diagonalisierbar: $AV = V\Lambda$, mit $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_m\}$

$$\stackrel{y=Vz}{\Rightarrow} z' = \Lambda z \iff z'_j = \lambda_j z_j$$

$$z(0) = z_0 := V^{-1}y_0 \quad z_j(0) = z_{0,j}$$

$$\Rightarrow z_j(t) = e^{\lambda_j t} \cdot z_{0,j} \Rightarrow z(t) = e^{\Lambda t} z_0$$

$$\stackrel{z=V^{-1}y}{\Rightarrow} y(t) = e^{At} y_0 = \sum c_i \cdot e^{\lambda_i t} \cdot \vec{z}_i$$

Definition: Der Begriff *steif* bedeutet hier, dass

$$\min_j \Re(\lambda_j) \ll \max_j \Re(\lambda_j) < 0.$$

Bemerkung: Die Situation ist qualitativ analog, wenn A nicht konstant oder sogar wenn $f(t, y)$ schwach nichtlinear in y ist.

$$A \approx \frac{\partial f}{\partial y} f(t, y)$$

Bei den besprochenen Methoden (RK, LMSV) ist das folgende Diagramm kommutativ:

$$\begin{array}{ccc} y' = Ay, y_0 & \xrightarrow{y=Vz} & z' = \Lambda z, z_0 \\ \text{num. Lösung} \downarrow & & \downarrow \text{num. Lösung} \\ \{y_n\}_{n=0}^\infty & \xrightarrow{y=Vz} & \{z_n\}_{n=0}^\infty \end{array}$$

Also: es genügt die skalare Differentialgleichung $y' = \lambda y$ zu betrachten, wobei $\lambda \in \mathbb{C}$.

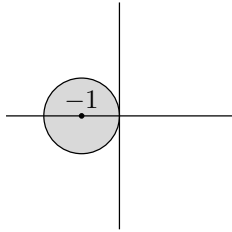
Oft kann man die entsprechende numerische Lösung explizit angeben.

Beispiel: Explizites Euler-Verfahren:

$$\begin{aligned} y_{n+1} &= y_n + h \underbrace{f(t_n, y_n)}_{\lambda y_n} \\ &= (1 + h\lambda)y_n = (1 + h\lambda)^{n+1}y_0 \end{aligned}$$

Es gilt:

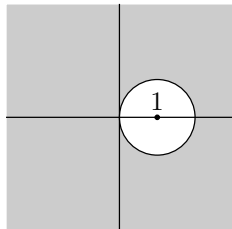
$$\begin{aligned} y_n \rightarrow 0 &\iff |1 + h\lambda| < 1 \\ \{y_n\} \text{ beschränkt} &\iff |1 + h\lambda| \leq 1 \end{aligned}$$



Beispiel: Implizites Euler-Verfahren:

$$\begin{aligned} y_{n+1} &= y_n + h \underbrace{f(t_{n+1}, y_{n+1})}_{\lambda y_{n+1}} = y_n + h\lambda y_{n+1} \\ \Rightarrow y_{n+1} &= \frac{1}{1 - h\lambda} y_n \\ &= \frac{1}{(1 - h\lambda)^{n+1}} y_0 \end{aligned}$$

$$\begin{aligned} y_n \rightarrow 0 &\iff |1 - h\lambda| > 1 \\ \{y_n\} \text{ beschränkt} &\iff |1 - h\lambda| \geq 1 \end{aligned}$$



Das heisst: Falls $y(t) \rightarrow 0$ für $t \rightarrow \infty \Rightarrow \Re(\lambda) < 0 \xRightarrow{\text{impl. Euler}} y_n \rightarrow 0$, das heisst, wenn die exakte Lösung gedämpft ist, so ist es auch die numerische.

Definition: Ein *Stabilitätsgebiet* \mathcal{S} eines Verfahrens ist eine Menge

$$\mathcal{S} := \left\{ z = h\lambda \in \mathbb{C} \mid \begin{array}{l} \text{numerische Lösung } \{y_n\}_{n \geq 0} \text{ zu } y' = \lambda y \\ \text{ist beschränkt für beliebige Anfangswerte} \end{array} \right\} \quad (5.2)$$

Bemerkung: Bei den LMSV brauchen die Anfangswerte nicht auf einer Lösungskurve zu liegen.

Definition: Dahlquist 1963:

Ein Verfahren heisst *A-stabil*, falls die numerische Lösung $(y_n)_{n \geq 0}$ beschränkt ist für jede Differentialgleichung $y' = \lambda y$ mit $\Re(\lambda) \leq 0$ und für beliebige Anfangswerte.

Satz: Ein Verfahren ist A-stabil genau dann, wenn $\{z \in \mathbb{C} \mid \Re(z) \leq 0\} \subset \mathcal{S}$ (Beweis: trivial; Umformulierung der Definition)

Beispiel: Das implizite Euler-Verfahren ist A-stabil; das explizite Eulerverfahren nicht.

Satz: LMSV stabil (0-stabil) $\iff 0 \in \mathcal{S}$
(Beweis: trivial; Umformulierung der Definition)

Beispiel: Trapez-Regel:

$$y_{n+1} = y_n + \frac{h}{2}(f(t_n, y_n) + f(t_{n+1}, y_{n+1}))$$

(Nebenbemerkung: Dieses Verfahren ist ein RKV und ein LMSV (eigentlich ein implizites Einschrittverfahren)).

Aus der Funktionentheorie ist bekannt, dass

$$\varphi(z) = \frac{1+z}{1-z}, \Re(z) \leq 0 \Rightarrow |\varphi(z)| \leq 1$$

$$\stackrel{f=\lambda y}{\Rightarrow} y_{n+1} = y_n + \frac{h\lambda}{2}(y_n + y_{n+1})$$

$$y_{n+1} = \frac{1 + \frac{h\lambda}{2}}{1 - \frac{h\lambda}{2}} y_n$$

$$\{y_n\}_{n \geq 0} \text{ beschränkt} \iff \underbrace{\left| \frac{1 + \frac{h\lambda}{2}}{1 - \frac{h\lambda}{2}} \right|}_{\text{Möbius-Transformation}} \leq 1$$

$$\Re(h\lambda) \leq 0 \xrightarrow{\text{M-T}} \left| \frac{1 + \frac{h\lambda}{2}}{1 - \frac{h\lambda}{2}} \right| \leq 1$$

\Rightarrow Stabilitätsgebiet \mathcal{S} = linke Halbebene.

\Rightarrow A-stabil

(In gewissem Sinne optimal; \mathcal{S} so klein wie möglich für A-stabiles Verfahren.)

Beispiel: Stabilitätsgebiete expliziter RK-Verfahren

Wenden RKV an auf $y' = \lambda y$:

$$\begin{cases} \mathbf{y}_i = y_n + h\lambda \sum_{j=1}^{i-1} a_{ij} \mathbf{y}_j, & i = 1, \dots, s \\ y_{n+1} = y_n + h\lambda \sum_{j=1}^s b_j \mathbf{y}_j \end{cases} \quad (5.3)$$

Mit Induktion folgt:

Lemma:

$$\begin{aligned} y_i &= p_i(h\lambda) y_n, & p_i &\in \mathcal{P}_{i-1} \\ y_{n+1} &= p_0(h\lambda) y_n, & p_0 &\in \mathcal{P}_s \setminus \mathcal{P}_0 \end{aligned}$$

($p_0 \notin \mathcal{P}_0$, weil nicht alle $a_{ij}, b_j = 0$.)

Satz: Das Stabilitätsgebiet $\{h\lambda = z \in \mathbb{C} \mid |p_0(z)| \leq 1\}$ eines expliziten RKV ist beschränkt.

Beweis: Es ist $\mathcal{S} = \{z \in \mathbb{C} \mid |p_0(z)| \leq 1\}$.

Da $p_0 \in \mathcal{P}_s \setminus \mathcal{P}_0$ (also $(p_0(z) \rightarrow \infty \text{ für } z \rightarrow \infty)$), ist \mathcal{S} beschränkt. \square

Korollar: Explizite RKV sind nicht A-stabil.

Stabilitätsgebiete von LMSV: Wenden LMSV auf $y' = \lambda y$ an:

$$\sum_{j=0}^k \alpha_j y_{n-j} = h\lambda \sum_{j=0}^k \beta_j y_{n-j} \quad (5.4)$$

das heisst

$$\sum_{j=0}^k (\alpha_j - h\lambda\beta_j) y_{n-j} = 0 \quad (5.4')$$

Homogene lineare Differentialgleichung mit konstanten Koeffizienten mit charakteristischem Polynom

$$\sum_{j=0}^k (\alpha_j - h\lambda\beta_j) \zeta^{k-j} = \rho(\zeta) - h\lambda\sigma(\zeta) \quad (5.5)$$

Definition: Annahme: ρ und σ sind teilerfremd; ein solches LMSV heisst *irreduzibel*

Satz: Das Stabilitätsgebiet eines irreduziblen LMSV ist gegeben durch

$$\mathcal{S} = \{z = h\lambda \in \mathbb{C} \mid \rho(\zeta) - h\lambda\sigma(\zeta) = 0 \Rightarrow |\zeta| \leq 1, |\zeta| < 1 \text{ falls mehrfache NS}\} \quad (5.6)$$

Beweis: Analog zum Beweis der 0-Stabilität. □

Satz: Ein irreduzibles LMSV ist A-stabil genau dann, wenn

$$|\zeta| > 1 \Rightarrow \Re\left(\frac{\rho(\zeta)}{\sigma(\zeta)}\right) > 0 \quad (5.8)$$

Beweis:

1. Es gelte (5.8) und es sei ζ_0 Nullstelle von $\rho(\zeta) - h\lambda\sigma(\zeta)$ für ein $h\lambda = z_0$ und $\Re(z_0) \leq 0$.

Dann ist $\sigma(\zeta_0) \neq 0$ weil LMSV irreduzibel, also $z_0 = \rho\zeta$

zu zeigen bleibt $|\zeta_0| = 1 \Rightarrow$ einfache Nullstelle:

Aus (5.8) folgt wegen Stetigkeit von $\frac{\rho}{\sigma}$ in ζ_0 , dass $\Re(h\lambda) = \Re(z_0) \geq 0$

Kritisch ist nur $\Re(z_0) = 0$:

Nun hat $\frac{\rho}{\sigma}$ in ζ_0 eine Taylor-Entwicklung

$$\frac{\rho(\zeta)}{\sigma(\zeta)} - \underbrace{\frac{\rho(\zeta_0)}{\sigma(\zeta_0)}}_{=z_0} = c_1(\zeta - \zeta_0) + c_2(\zeta - \zeta_0)^2 + \dots$$

Aus (5.8) folgt, dass $c_1 \neq 0$ ist; andernfalls würde eine kleine Kreisscheibe $U_\varepsilon(\zeta_0)$ um ζ_0 durch $\frac{\rho}{\sigma}$ auf eine zweifach bedeckte Umgebung von z_0 abgebildet; es wäre also nicht für alle $|\zeta| > 1$ in $U_\varepsilon(\zeta_0)$ $\Re(\frac{\rho(\zeta)}{\sigma(\zeta)}) > 0$. □

Beispiel: Adams-Moulton:

$k = 0$: implizites Euler: A-stabil, $p = 1$

$k = 1$: Trapez-Regel: A-stabil, $p = 2$

$k = 2$: $\rho(\zeta) = \zeta^2 - \zeta$; $\sigma(\zeta) = \frac{5}{12}\zeta^2 + \frac{8}{12}\zeta - \frac{1}{12}$

Nullstellen von σ : $\zeta_{1,2} = -\frac{4}{5} \pm \frac{\sqrt{21}}{5}$

$|\zeta_2| > 1 \Rightarrow \mathcal{S}$ beschränkt.

$k > 2$: wie für $k = 2$.

Beispiel: BDF-Verfahren:

$$\rho(\zeta) = \zeta^k \sum_{j=1}^k \frac{1}{j} \left(1 - \frac{1}{\zeta}\right)^j; \quad \sigma(\zeta) = \zeta^k$$

$k = 1$: implizites Euler: A-stabil

$k = 2$: $\rho(\zeta) = \frac{3}{2}\zeta^2 - 2\zeta + \frac{1}{2}$; $\sigma(\zeta) = \zeta^2$

A-stabil (man zeigt: $\Re(\frac{\rho(\zeta)}{\sigma(\zeta)}) > 0$, falls $|\zeta| > 1$, siehe Serie 5, Aufgabe 3)

$3 \leq k \leq 6$: $A(\alpha)$ -stabil mit folgenden α :

k	1	2	3	4	5	6
α	90°	90°	86°	73°	51°	18°

$k > 6$: nicht A-stabil

Definition: Ein Verfahren heisst $A(\alpha)$ -stabil mit $0 < \alpha \leq \frac{\pi}{2}$, wenn \mathcal{S} den Sektor $\{z \in \mathbb{C} \mid |\arg(-z)| \leq \alpha\}$ enthält.

5.1 Ordnungsschranke für A-stabile LMSV

Satz: Dahlquist 1963:

1. Für die Ordnung eines A-stabilen LMSV gilt $p \leq 2$.
2. Ist $p = 2$, so gilt in $\rho(e^z) - z\sigma(e^z) = C_{p+1}z^{p+1} + \dots: C_3 \leq -\frac{1}{12}$
3. $p = 2$ und $C_3 = -\frac{1}{12}$ gilt nur für die Trapezregel

Index

Symbols

$A(\alpha)$ -stabil	26
0-stabil	20
1. charakteristische Polynom	20
2. charakteristische Polynom	20

A

A-stabil	23
Anfangswert	1

B

BDF-Verfahren	18
Butcher-Tableau	13

D

Differentialgleichung 1. Ordnung	1
Diskretisationsfehler	10

E

Einschrittverfahren	8
einseitige Lipschitz-Bedingung	3
ESV	8

F

Fehlerordnung	9
---------------------	---

I

implizites Euler-Verfahren	17
irreduzibel	25

K

Kollokationsbedingung	18
konsistent	20
Konsistenzbedingung	10
konvergent	21

L

Lösung	1
Lineare Mehrschritt-Verfahren	15
LMSV	15
explizites	15
implizites	15
lokale Fehler	8

O

Ordnung	
LMSV	18

R

Rundungsfehler	10
----------------------	----

S

Stabilitätsgebiet	23
steif	22

T

Taylorreihen-Verfahren	11
Trapezregel	17
truncation error	19