Numerik und Statistik: Numerik

Dimitri Ovrutskiy

29. Juni 2012

Inhaltsverzeichnis

1 Vorwort **3**

2 Einfuhrung **4**
2.1 Grundbegriffe, Definitionen 4
2.2 Beispiele, Modellierung 7
2.3 (Explizites) Eulersches Polygonzugverfahren 11

3 Existenzsätze; Euler-Verfahren **12**
3.1 Existenzsatz von Peano 12
3.2 Picard-Lindelöf Iteration 15
3.3 Taylorreihe der AWP-Lösung 15
3.4 Konvergenz des (expl.) Euler-Verfahrens 16
3.5 Systeme der Differentialgleichungen 17

4 Weitere Einschrittverfahren: Runge-Kutta Methoden **20**
4.1 Grundbegriffe und Definitionen, Butcher-Schema 20
4.2 Runge-Kutta Verfahren 2. Ordnung 25
4.3 Runge-Kutta Verfahren der Ordnung 3 und 4 26
4.4 Konvergenzresultate für die R.-K.-Methoden 30
4.5 Schrittweitensteuerung, eingebettete R.-K.-Methoden 31
4.6 Implizite R.-K.-Methoden (ein Ausblick) 32

5 Randwertprobleme **34**
5.1 Aufgabestellung, Beispiele, Existenzresultate 34
5.2 Einfach- und Mehrfachschießverfahren 37
5.2.1 Das Einfachschießverfahren ('Wasserschlauchmethode') . 37
5.2.2 Das Mehrfachschießverfahren ('Multiple Shooting Method', MSV) 40
5.3 Differenzenverfahren 43

6 Anhang **46**
6.1 Landau-Symbole 46
6.2 Eigenwerte und Eigenvektoren 49
6.3 Orthogonale Matrizen 52
6.4 Symmetrische Matrizen 53
6.5 Numerische Differentiation (Finite Differenzen) 54

1 Vorwort

Während des Numerik-Teils der Vorlesung 'Numerik und Statistik', die im 8. Semester für die auf Master Studierenden der Fachrichtung Mechttronik und Sensortechnik an der Hochschule der Technik und Wirtschaft des Saarlandes gehalten wird, werden theoretische und praktische Aspekte der numerischen Lösung (gewöhnlicher) Differentialgelichungen besprochen. Als Grundlage dienten die von Hr. Prof. Dr. A.Arnold in Saarbrücken und Münster und von Hr. Prof. Dr. S. Rjasanow in Saarbrücken gehaltene Vorlesungen 'Höhere Numerik 1', 'Praktische Mathematik 2', 'Gewöhnliche Differentialgleichungen'.

Bei dem FT-Studium haben praktische Aspekte und Anwendungen vor reiner Theorie Vorrang, deswegen wurden viele (auch interessante) Beweise weggelassen und stark auf Anwendbarkeit geachtet.

2 Einfuhrung

2.1 Grundbegriffe, Definitionen

Im weiteren seien $n \in \mathbb{N}$, $F : D \subset \mathbb{R}^{n+2} \to \mathbb{R}$, $f : G \subset \mathbb{R}^{n+1} \to \mathbb{R}$. D, G seien Gebiete, d.h. zusammenhängend und konvex.

Definition 1 *Eine Bestimmungsgleichung für $y = y(x)$ der Form*

$$F(x, y, y', y'', ..., y^{(n)}) = 0 \tag{1}$$

heißt implizite gewöhnliche Differentialgleichung (DGl) n-ter Ordnung (iDGL). Die Gleichung

$$y^{(n)} = f(x, y, y', ..., y^{(n-1)}) \tag{2}$$

heißt explizite DGL n-ter Ordnung (eDGL).

Eine n-mal stetig differenzierbare Funktion $y : I \to \mathbb{R}$ ($y \in C^n(I)$) heißt (klassische) Lösung von iDGL (1) auf dem offenen Intervall $I \subset \mathbb{R}$, wenn $\forall x \in I$ gilt

$$(x, y(x), y'(x), ..., y^{(n)}(x)) \in D$$

und

$$F(x, y(x), y'(x), ..., y^{(n)}(x)) = 0.$$

Analog fur (2):

$$(x, y(x), y'(x), ..., y^{(n-1)}(x)) \in G$$

und

$$f(x, y(x), y'(x), ..., y^{(n)}(x)) = y^{(n)}(x).$$

$y' = f(x, y)$ ist die explizite DGL 1. Ordnung.

Die Lösung von DGL ist gewönlich ein Kurvenschar, d.h. die Lösung ist nicht eindeutig und enthält sogennante Integrationskonstanten.

Definition 2 *Ein Anfangswertproblem (AWP) ist eine DGL*

$$y^{(n)} = f(x, y, y', ..., y^{(n-1)})$$

mit vorgegebenen Werten

$$y(x_0) = y_0, y'(x_0) = y_1, ..., y^{(n-1)}(x_0) = y_{n-1},$$

$$(x_0, y_0, ..., y_{n-1}) \in G.$$

Das AWP heißt lokal lösbar, wenn $\exists \epsilon > 0 : \ \forall x \in I = (x_0 - \epsilon, x_0 + \epsilon) \ \exists y = y(x)$, die die DGL mit den Anfangsbedingungen (AB) erfüllt.

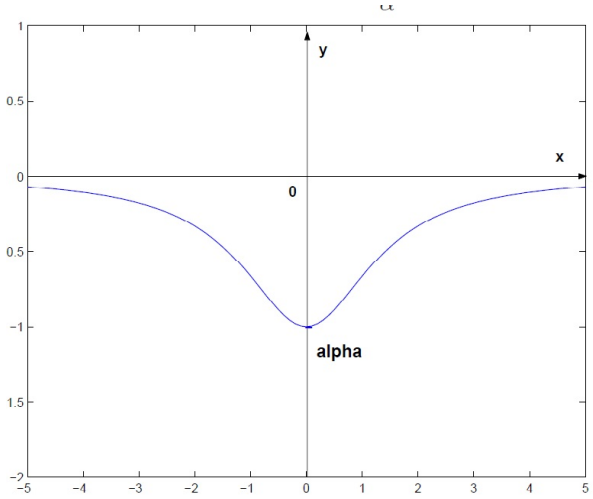
Beispiel 3 $y' = xy^2,\ y(0) = \alpha \in \mathbb{R}$

$\Rightarrow$ *DGL mit trennbaren Variablen*

$$\Rightarrow \frac{dy}{y^2} = x\,dx, \qquad (y \neq 0 \text{ !!!})$$

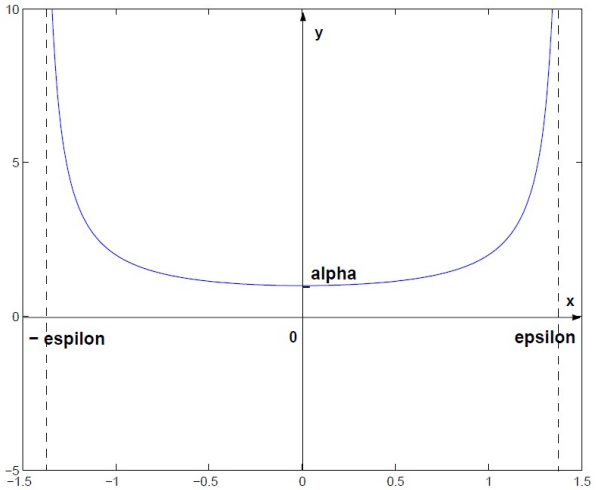
- $\rightarrow -\frac{1}{y} = \frac{x^2}{2} + \tilde{c} \qquad \Rightarrow \quad y = \frac{-2}{x^2+c}$
- $y = 0 \quad \Rightarrow$ *spezielle Lösung*
- *AWP:* $y(0) = -\frac{2}{c} = \alpha \quad \Rightarrow \ c = -\frac{2}{\alpha}, \quad (\alpha \neq 0 \text{ !!!})$

  1. $\alpha = 0 \ \Rightarrow \ y \equiv 0$ *(global)*
  2. $\alpha < 0 \ \Rightarrow \ y = \frac{-2}{x^2 - \frac{2}{\alpha}}$ *(globale Lösung)*



$$y = \frac{-2}{x^2+1}$$

3. $\alpha > 0,\ y(x) = \frac{-2}{x^2 - \frac{2}{\alpha}}$ *(lokale Lösung für* $|x| < \epsilon = \sqrt{\frac{2}{\alpha}}$*)*



Setzt man

$$y_1(x) := y(x),$$
$$y_2(x) := y'(x),$$
$$.....$$
$$y_n(x) := y^{(n-1)}(x)$$

in (2) ein, so laßt sich das AWP n-ter Ordnung in ein DGL-System 1. Ordnung umwandeln:

$$\left\{ \begin{array}{lcll} y_1'(x) & = & y_2(x) & =: \quad f_1(x, y_1, ..., y_n) \\ y_2'(x) & = & y_3(x) & =: \quad f_2(x, y_1, ..., y_n) \\ & & ..... & \\ y_n'(x) & = & f(x, y_1(x), ..., y_n(x) & =: \quad f_n(x, y_1, ..., y_n) \end{array} \right. ,$$

$$y_i(x) = y_i, \quad i = 1, ..., n,$$

oder, in Vektorform

$$\left\{ \begin{array}{l} y'(x) = f(x, y), \quad y = (y_1(x), ..., y_n(x))^T \in \mathbb{R}^n, \ f = (f_1, ..., f_n)^T) \\ y(x_0) = y_0. \end{array} \right.$$

Eine 'Rückumwandlung' solcher Vektorsystem in eine skalare Gleichung n-ter Ordnung geht i.A. nicht.

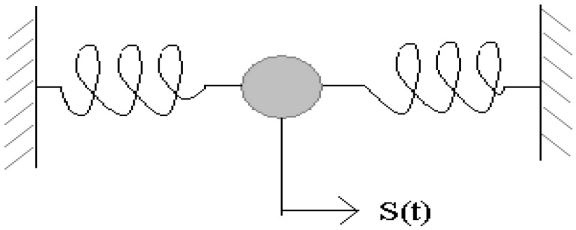
D.h., wir betracheten immer Systeme 1. Ordnung.

Definition 4 *(Hadamard, 1865-1963) Das AWP heißt korrekt gestellt, wenn drei Eigenschaften gesichert sind:*

1. *Die Existenz einer lokalen Lösung*
2. *Die Eindeutigkeit der lokalen Lösung*
3. *Die stetige Abhängigkeit der lokalen Lösung von den AB (Stabilität)*

2.2 Beispiele, Modellierung

Beispiel 5 *"Lineare Pendelbewegung"*



$s(t)$ - *Auslenkung;*
$\dot{s}(t)$ - *Geschwindigkeit;*
$\ddot{s}(t)$ - *Beschleunigung.*

- *2. Newton'sches Gesetz:*

$$ma = F \quad (m\ddot{s} = F)$$

- *Hooke'sches Gesetz:*

$$F = -ks(t)$$

Mit $\omega^2 := \frac{k}{m}$

$$\ddot{s}(t) + \omega^2 s(t) = 0.$$

Als System:

$$\begin{cases} \dot{s}(t) &= v(t) \\ \dot{v}(t) &= -\omega^2 s(t) \end{cases}$$

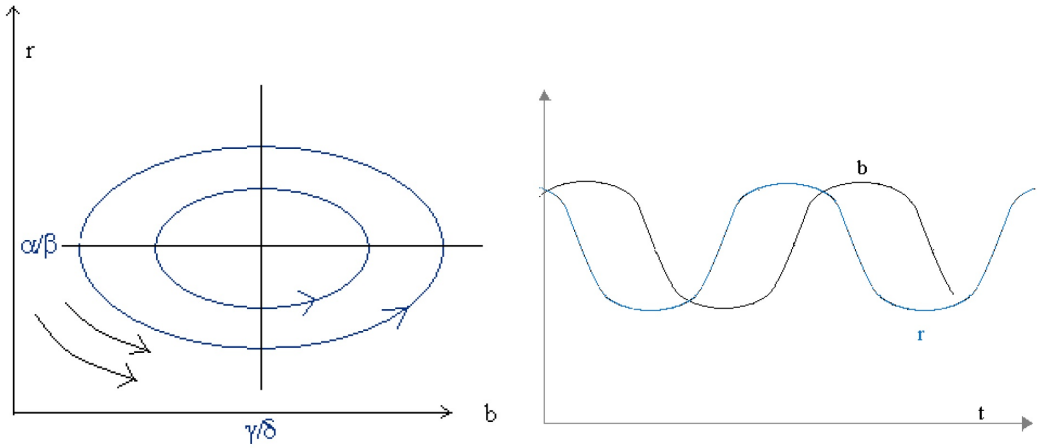
Anfangsbedinugnen (AB):
$s(0) = s_0$
$v(0) = v_0$.
In Matrixform (lineares Gleichungssystem):

$$\begin{pmatrix} \dot{s} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} \begin{pmatrix} s \\ v \end{pmatrix}$$

Beispiel 6 *"Räuber-Beute Systeme"*
Z.B.: Hase-Luchs für $b(t)$ *und* $r(t)$ *entsprechend:*

$$\begin{cases} \dot{b} &= \alpha b - \beta br \\ \dot{r} &= -\gamma r + \delta br \end{cases}$$

mit positiven Konstanten $\alpha,\ \beta,\ \gamma,\ \delta > 0$.



Ruhelage, geschlossene Trajektorien.

Beispiel 7 *"Turbulenzen in der Erdathmosphäre"*

Lorenz-System ist eine 3D-Moden Galerkin-Approximation der Boussinesq-Gleichung, die die Konvektion einer von unten beheitzten 2D-Schicht beschreibt.

Sein chaotisches Verhalten weckt die Gedanken über die Einschränkungen in Wettervorhersagen.

$x(t)$ *- konvektive Luftbewegung / Geschwindigkeit*

$y(t)$ *- horizontale Temperaturveränderung*

$z(t)$ *- vertikale Temperaturveränderung*

a *und* b *sind Zahlen, die den speziellen physikalischen (meteorologischen) Konstanten entsprechen:*

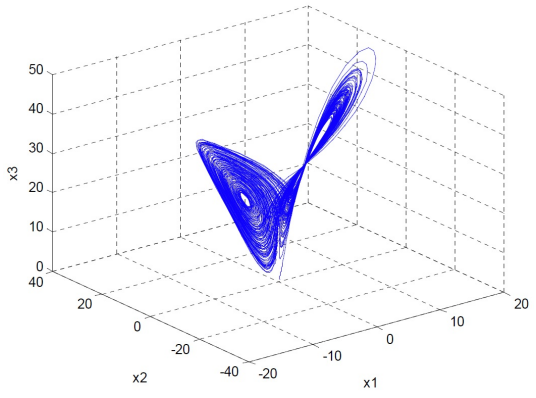
a *Prandtzahl,*

$(1 + b^2)$ *Rayleighzahl.*

Dann lautet das Lorenz-System

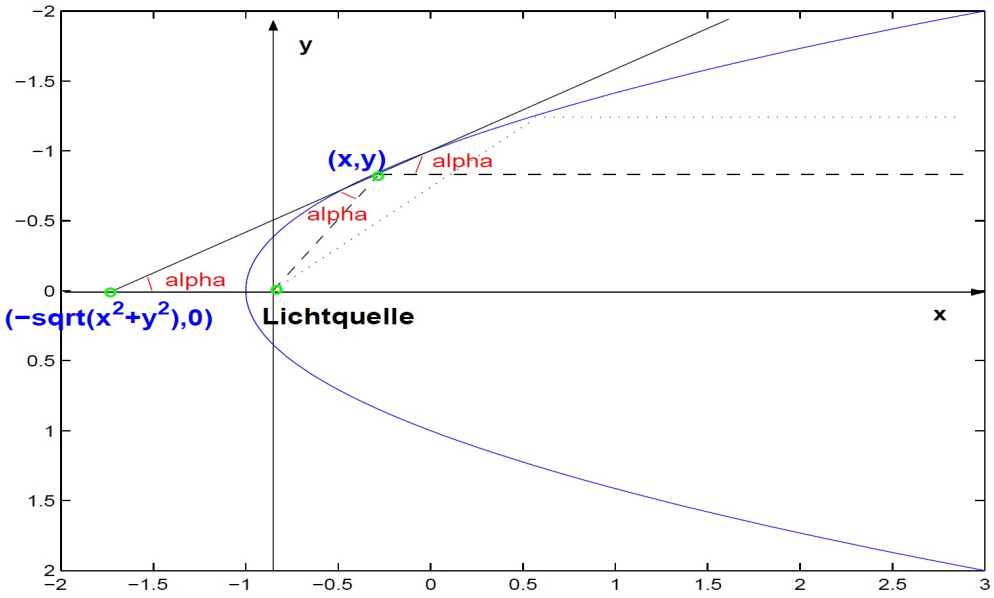
$$\begin{cases} \dot{x}(t) &= a(y - x) \\ \dot{y}(t) &= (1 + b^2)x - y - xz \\ \dot{z}(t) &= xy - z \end{cases}$$

$\Rightarrow$ *Lorenz-Attraktor (1963, E.N. Lorenz)*



Beispiel 8 *"Scheinwerfer"*

Gesucht ist die Meridiankurve einer Rotationsfläche (Drehachse ist die x-Achse), die alle vom Ursprung ausgehenden Strahlen in Richtung der x-Achse reflektiert (alle gespiegelten Strahlen sind dann parallel zur x-Achse).



Wegen Heron'sches Spiegelungsgesetzt (Heron'sche Katoptrik) und den Satz über die Winkel zwischen einer Sekanten und den parallelen Geraden, sowie wegen der Tatsache, daß Anstieg einer Tangenten dem Ableitungswert der Funktion im Berührungspunkt gleich ist, erhalten wir in jedem Punkt (x, y) der gesuchten Kurve (mit Anwendung der geometrischen Bedeutung von tan *in einem rechtswinkligen Dreieck)*

$$\Rightarrow \begin{cases} \tan\alpha = y' \\ \tan\alpha = \frac{y-0}{x+\sqrt{x^2+y^2}} \end{cases}$$

$$\Rightarrow y' = \frac{y}{x+\sqrt{x^2+y^2}}$$

Diese DGl läßt sich wie folgt umformen (s. Mathematik 3 bzw. Paupla):

$$y' = \frac{y}{x+\sqrt{x^2+y^2}} = \frac{\frac{y}{x}}{1+\sqrt{1^2+\left(\frac{y}{x}\right)^2}} = f\left(\frac{y}{x}\right)$$

Substitution $v = \frac{y}{x} \Rightarrow v'x + v = y'$ *führt auf eine DGl mit trennbaren Variablen:*

$$v'x + v = \frac{v}{1+\sqrt{1+v^2}}$$

$$\Rightarrow \frac{1+\sqrt{1+v^2}}{v\sqrt{1+v^2}}dv = -\frac{dx}{x}$$

$$\rightarrow \ln \frac{v^2}{1+\sqrt{1+v^2}} = -\ln x + \ln c = \ln \frac{c}{x}$$

$$\Rightarrow v^2 \frac{x}{c} - 1 = \sqrt{1+v^2}$$

$$\Rightarrow \frac{v^2 x^2}{c^2} = 1 + \frac{2x}{c}$$

Nach der Rücksubstitution erhalten wir

$$y^2 = c^2 + 2cx,\ c \in \mathbb{R},$$

eine konfokale Schar von Parabeln.

2.3 (Explizites) Eulersches Polygonzugverfahren

Für das AWP $y' = f(x,y),\ y(x_0) = y_0,\ y : \mathbb{R} \to \mathbb{R}^n$ existiert i.A. kein explizites Lösungsverfahren $\Rightarrow$ numerische Näherungen für $y(x),\ x > x_0$.

- *Diskretisierung des Intervals* $[x_0, X]$ durch die äquidistanten Knoten

$$x_k = x_0 + k \cdot h,\ h = \frac{X - x_0}{n},\ k = 0, 1, ..., n.$$

  Berechnung der Näherungen

$$y_0, y_1, ..., y_n$$

  für

$$y(x_0), y(x_1), ..., y(x_n)$$

  durch

$$y_{k+1} = y_k + h \cdot f(x_k, y_k), \quad k = 0, 1, ..., n-1.$$

  Motivation:

$$y)x_{k+1} - y(x_k) = \int_{x_k}^{x_{k+1}} f(x,y)\mathrm{dx} \overset{Rechtecksregel}{\approx} hf(x_k, y(x_k))) \approx hf(x_k, y_k).$$

- *Ungleichmäßige Diskretisierung:*

$$x_0 < x_1 < ... < x_n = X.$$

  Bezeichnungen:

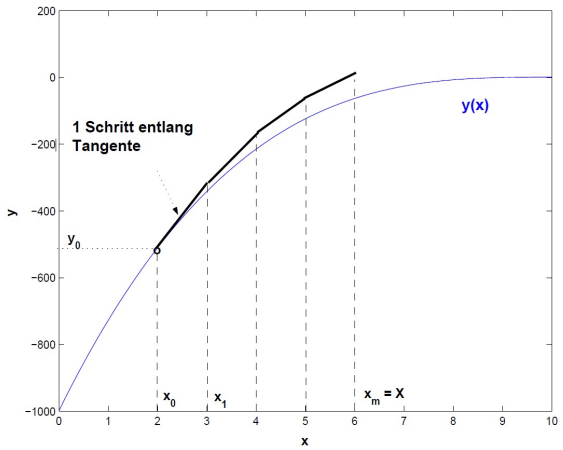
$$h_i = x_{i+1} - x_i, \quad i = 0, 1, ..., n-1,$$

$$h = \max_{0 \le i \le n-1} h_i.$$

- *Polygonzug:*

$$y_h(x) = y_i + (x - x_i) f(x_i, y_i), \quad x_i \le x \le x_{i+1}, \quad x \in [x_0, X].$$

  $y_h(x)$ ist eine stückweise lineare Funktion; Näherung für y(x).



Frage: Was passiert mit $y_h(x)$ für $h \to 0$?

3 Existenzsätze; Euler-Verfahren

3.1 Existenzsatz von Peano

Sei $D = \{(x, y) \in \mathbb{R}^2 : |x - x_0| \leq a, |y - y_0| \leq b\}$.

Satz 9 *(G.Peano, 1858-1932) Sei $f(x, y) : D \to \mathbb{R}$ stetig. Dann besitzt das AWP*

$$y' = f(x, y), \quad y(x_0) = y_0$$

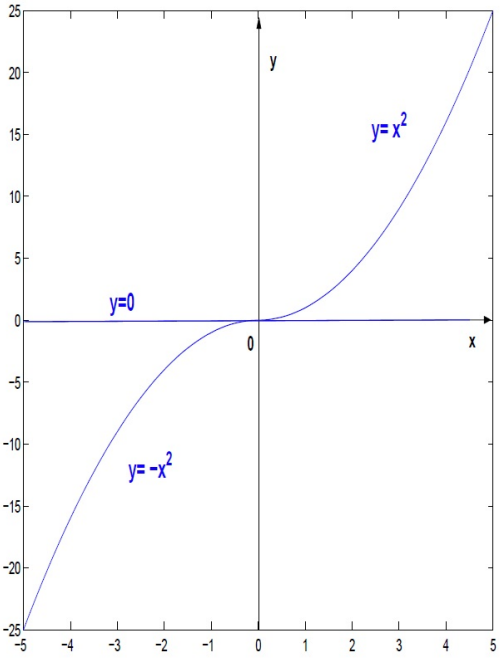
mindestens eine Lösungskurve, die sich bis zum Rand von D erstreckt.

Bemerkung 10 *Die Steitgkeit von $f(x, y)$ alleiche reicht nicht aus, um die Eindeutigkeit der Lösung zu sichern.*

Beispiel 11

$$y' = 2\sqrt{|y|}, \quad y(0) = 0$$

hat zwei stetig-differenzierbare Lösungen: a) $y \equiv 0$ b) $y = sgn(x) \cdot x^2$



Bemerkung 12 *$y(x)$ braucht nicht auf dem ganzen Intervall $|x - x_0| \leq a$ zu existieren (Beispiel 3).*

Definition 13 *Sei $f : G \subset \mathbb{R}^2 \to \mathbb{R}$, G Gebiet.*

1. *$f(x, y)$ erfüllt auf G eine Lipschitz-Bedingung bezüglich y (R, Lipschitz, 1832-1903), wenn eine (globale) Konstante $L \geq 0$ existiert , so daß $forall(x, z), (x, y) \in G$*

$$|f(x, z) - f(x, y)| \leq L|z - y|.$$

2. *$f(x,y)$ erfüllt auf G eine lokale Lipschitz-Bedingung bezüglich y, wenn es zu jedem Punkt von G eine Umgebung U gibt, in der f eine Lipschitz-Bediungung bzgl y erfüllt, d.h. $L = L(U)$.*

Beispiel 14 1. *$f(x,y,) = |y| \Rightarrow L = 1$, globale L-Bedingung.*

2. *$f(x,y) = y^2 \Rightarrow \frac{|f(x,z)-f(x,y)|}{|z-y|} = |z+y| \to$ lokale L-Bedingung, aber nicht globale*

3. *$f(x,y) = 2\sqrt{|y|}$ (s. Beispiel 3) für $x = 0 \Rightarrow \frac{|f(0,0)-f(0,y)|}{|y|} = \frac{2}{\sqrt{|y|}} \overset{y\to 0}{\to} \infty$ $\Rightarrow$ keine L-Bedingung!*

Satz 15 *Die Funktion $f(x,y)$ sei stetig und erfülle auf dem Gebiet $G \subset \mathbb{R}^2$ eine L-Bedingung bezüglich y. Die Lösung des AWP's*

$$y' = f(x,y), \quad y(x_0) = y_0$$

ist eindeutig.

In Kombination mit dem Satz (9) (Peano) ergibt sich dann

Satz 16 *(Picard)*
Die Funktion $f(x,y)$ sei auf dem Gebiet $G \subset \mathbb{R}^2$ stetig und erfülle dort eine lokale L-Bedingung bezüglich y. Dann gibt es zu jedem $(x_0,y_0) \in G$ für das AWP

$$y' = f(x,y), \quad y(x_0) = y_0$$

eine eindeutige Lösungskurve $y = y(x)$, die sich bis an den Rand von G erstreckt.

Aus Lipschitz-Bedingung an f folgt Stabilität der Differentialgleichung:

Satz 17 *(Stabilität bzgl. AB)*
Erfüllt die stetige Funktion $f(x,y)$ im Gebiet $G \subset \mathbb{R}^2$ eine L-Bedingung, dann gilt für

$$y_1' = f(x,y_1), \quad y_1(x_0) = y_0^{(1)}$$
$$y_2' = f(x,y_2), \quad y_2(x_0) = y_0^{(2)}$$

die Abschätzung

$$|y_1(x) - y_2(x)| \le e^{L(x-x_0)}|y_0^{(1)} - y_0^{(2)}|.$$

Beispiel 18 *(inhärente Instabilität / Sensitivität)*
Sei

$$A = \begin{pmatrix} 0 & 1 \\ 10 & 9 \end{pmatrix}, \qquad y(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix},$$
$$y' = Ay, \qquad y(0) = (\alpha,\ -\alpha)^T$$

das AWP.

- *$f(y) = Ay$ ist linear, daher Lipschitz-stetig $\Rightarrow$ das AWP ist stabil bzgl. AB.*
- *allgemiene Lösung von $y' = Ay$ lautet*

$$y(x) = c_1 e^{-x} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + c_2 e^{10x} \begin{pmatrix} 1 \\ 10 \end{pmatrix}.$$

  λ_1 und λ_2 unterscheiden sich stark!

- $\begin{pmatrix} \alpha \\ -\alpha \end{pmatrix}$ *liegt im Eigenraum von $\lambda_1 < 0$*

$\Rightarrow$ *kleine Änderungen der AB können große Änderungen in der Lösung verursachen*

$\Rightarrow$ *numerische Probleme!*

Z.B.:

$$y_\epsilon(0) := \begin{pmatrix} \alpha + \epsilon \\ -\alpha \end{pmatrix}, \quad \epsilon > 0$$

$$\Rightarrow \quad y_\epsilon(x) = \left(\alpha + \frac{10}{11}\epsilon\right) e^{-x} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \underbrace{\frac{\epsilon}{11} e^{10x} \begin{pmatrix} 1 \\ 10 \end{pmatrix}}_{Störung}$$

für $\alpha = 1,\ \epsilon = 10^{-8}$ $\Rightarrow$ relativer Fehler von $y_1(x = 2.5) \approx 800$.

Satz 19 *(Stabilität bzgl. der rechten Seite)*

Erfüllt eine stetige Funktion $f_1(x, y)$ auf $G \subset \mathbb{R}^2$ eine L-Bedingung bzgl. y und gelte für $f_2(x, y)$

$$|f_1(x, y) - f_2(x, y)| \leq \epsilon \quad \forall (x, y) \in G,$$

dann gilt fur die Lösungen

$$y_1' = f_1(x, y_1), \quad y_1(x_0) = y_0$$

$$y_2' = f_2(x, y_2), \quad y_2(x_0) = y_0$$

im Intervall $[x_0, X]$ die Abschätzung

$$|y_1(x) - y_2(x)| \leq \epsilon (X - x_0) e^{L(x - x_0)}.$$

Bemerkung 20 *Zusätzliche Stetigkeit und L-Bedingung an f_2 wäre nur für Existenz und eindeutigkeit von $y_2(x)$ nötig.*

3.2 Picard-Lindelöf Iteration

(E.Picard (1856-1941), E- Lindelöf (1870-1946))

Picard-Lindelöf Iteration: Explizite Methode zur Approxiamtion von Lösungen des AWP's

$$y' = f(x, y), \quad y_(x_0) = y_0.$$

Startfunktion:

$$y(x_0) \equiv y_0 \equiv const$$

Iteration:

$$y_{k+1} = y_0 + \int_{x_0}^{x} f(t, y_k(t))\mathrm{dt}, \qquad k = 0, 1, ...$$

Motivation: Umwandlung der DGL in eine Intagralgleichung durch integration beider Seiten unter der AB $y(x_0) = y_0$, die der Bestimmung der Integrationskonstante dient.

Satz 21 *Die Funktion $f(x, y)$ sein auf G stetig und erfülle eine L-Bedingung bzgl. y. Dann konvergiert die Funktionenfolge der P.-L.-Iteration auf einem Intervall $[x_0, x_0 + \epsilon]$, $\epsilon > 0$ gleichmäßig zur Lösung des AWP's.*

3.3 Taylorreihe der AWP-Lösung

AWP: $y'(x) = f(x, y), \quad y(x_0) = y_0$

Idee: Für die Berechnung von $y(x_0 + h)$ benutze einige Glieder der Taylor-Reihe:

$$y(x_0 + h) \approx y(x_0) + \frac{y'(x_0)}{1!}h + \frac{y''(x_0)}{2!}h^2 + ... + \frac{y^k(x_0)}{k!}h^k$$

1. Für $k = 1$ erhält man das Euler'sche Polygonzugverfahren.
2. Aus $y' = f(x, y)$ unmittelbar folgt

$$y'' = f_x(x, y) + f_y \cdot y' \quad \Rightarrow \quad y''(x_0) = f_x(x_0, y_0) + f_y(x_0, y_0) \cdot f(x_0, y_0);$$

$$y''' = f_{xx} + 2f_{xy}f + f_{yy}f^2 + f_y(f_x + f_y f)) \quad \Rightarrow ... \; u.s.w.$$

3. Mit Hilfe von Computer durchführbar (symbolische Formelmanipulationen)
4. Falls $f(x, y)$ analytisch in einer Umgebung von x_0, y_0, so konvergiert die Taylorreihe für $h \leq h_0$.

3.4 Konvergenz des (expl.) Euler-Verfahrens

AWP: $y' = f(x,y),\ y(x_0) = y_0$ Gesucht: $y(x),\ x \geq x_0$

$y_h(x) := y_i + (x - x_i) f(x_i, y_i),\ i = 0, ..., n-1;\ \ x_n = X;\ x \in [x_i, x_{i+1}]$ $(stückweise\, linear)$

Sei $G = \{(x,y) \in \mathbb{R}^2,\ x_0 \leq x \leq X,\ |y - y_0| \leq b\} = [x_0, X] \times [y_0 - b, y_0 + b]$.
Dann gilt

Satz 22 *(Stabilität des Euler-Verfahrens)*
Sei $x_0, x_1, ..., x_n = X$ eine feste Diskretisiereung des Intervalls $[x_0, X]$ und $y_h(x),\ z_h(x)$ die Eulerschen Polygonzüge für

$$y' = f(x,y), \quad y(x_0) = y_0,$$

$$z' = f(x,z), \quad z(x_0) = z_0.$$

Wenn die Funktion $f(x,y)$ auf G Lipschitz-stetig ist (Def. 13), dann gilt $\forall x \in [x_0, X]$ die Abschätzung

$$|y_h(x) - z_h(x)| \leq e^{L(x - x_0)} |y_0 - z_0|.$$

Beweis:
Es gilt

$$z_1 - z_0 = f(x_0, z_0)(x_1 - x_0)$$
$$y_1 - y_0 = f(x_0, y_0)(x_1 - x_0)$$

und damit

$$|z_1 - y_1| = |z_0 - y_0 + (x_1 - x_0)[f(x_0, z_0) - f(x_0, y_0)]| \leq$$
$$\leq |z_0 - y_0| + (x_1 - x_0)|f(x_0, z_0) - f(x_0, y_0)| \leq$$
$$\leq |z_0 - y_0| + (x_1 - x_0) L |z_0 - y_0| =$$
$$= (1 + L(x_1 - x_0))|z_0 - y_0| \leq$$
$$\leq e^{L(x_1 - x_0)} |y_0 - z_0|.$$

Für $|z_2 - y_2|$ gilt analog:

$$|z_2 - y_2| \leq e^{L(x_2 - x_1)} |y_1 - z_1 \leq e^{L(x_2 - x_1 + x_1 - x_0)} |y_0 - z_0|| = e^{L(x_2 - x_0)} |y_0 - z_0|,\ u.s.w.$$

$$\Rightarrow |y_h(x) - z_h(x)| \leq e^{L(x - x_0)} |y_0 - z_0| \ \ \forall x \in [x_0, X], \qquad q.e.d.$$

Für $h = max_{0 \leq i \leq n-1} h_i \to 0$ konvergieren die euler'sche Polyginzuge $y_h(x)$ gegen die Lösung des AWPs:

Satz 23 *Die Funktion $f(x,y)$ sei auf dem Gebiet*

$$G = \{(x,y) \in \mathbb{R}^2 | x_0 \leq x \leq X, |y - y_0| \leq b\}$$

stetig, $|f(x,y)| \leq A$ und erfülle auf G eine L-Bedingung bzgl. y. Dann gilt für $X - x_0 \leq \frac{b}{A}$:

1. *Die Euler'sche Polygonzüge $y_h(x)$ konvergieren für $h \to 0$ gleichmäßigegen eine stetige Funktion $\phi(x)$*

2. *$\phi(x)$ ist stetig differenzierbar und erfüllt die Gleichung $\phi' = f(x,y),\ \phi(x_0) = y_0,\ x_0 \le x \le X$ (d.h. ϕ ist die eindeutige Lösung)*

Unter zusätzlichen Voraussetzungen erhalten wir sogar eine Konvergenzabschätzung für das Euler'sche Polygonzugverfahren:

Satz 24 *Die Funktion $f(x,y)$ sei auf G stetig prtiell differenzierbar mit*

$$|f_x| \le M, \quad |f_y| \le L, \quad |f| \le A.$$

Dann kovergiert das Euler'sche Polygonzugverfahren, und der Fehler $y(x) - y_h(x)$ genügt der Abschätzung

$$|y(x) - y_h(x)| \le \frac{M + AL}{L}\left(e^{L(x-x_0)} - 1\right) h = O(h).$$

3.5 Systeme der Differentialgleichungen

Definition 25 *(Norm)*
Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ bzw. $\mathbb{K} = \mathbb{C}$. Eine Abbildung $||\cdot|| : V \to \mathbb{R}^+,\ x \mapsto ||x||$ heißt Norm von $x \in V$, falls sie folgende Eigenschaften hat (Normaxiomen):
$\forall x, y \in V,\ \forall \lambda \in \mathbb{K}$*:*

1. $||x|| = 0 \Leftrightarrow x = 0$ *(Definitheit)*

2. $||\lambda \cdot x|| = |\lambda| \cdot ||x||$ *(absolute Homogenität, Linearität)*

3. $||x + y|| \le ||x|| + ||y||$ *(Subadditivität, Dreiecksungleichung)*

Bemerkung 26 *(Eigenschaften einer Norm)*

1. $x = 0 \Rightarrow ||x|| = 0$

2. $||-x|| = ||x||$

3. $||x|| \ge 0$

4. $|\ ||x|| - ||y||\ | \le ||x - y||$

Bemerkung 27 *(Induzierte Normen)*
Eine Norm kann, muss aber nicht notwendigerweise, von einem Skalarprodukt $\langle \cdot, \cdot \rangle$ abgeleitet werden. Die Norm eines Vektors $x \in V$ ist dann definiert als

$$\|x\| := \sqrt{\langle x, x \rangle},$$

also die Wurzel aus dem Skalarprodukt des Vektors mit sich selbst. Man spricht in diesem Fall von der durch das Skalarprodukt induzierten Norm oder Hilbertnorm. Jede durch ein Skalarprodukt induzierte Norm erfullt die Cauchy-Schwarzsche Ungleichung

$$|\langle x, y \rangle| \le \|x\| \cdot \|y\|.$$

Sei uns das DGL-System im Vektorform gegeben

$$y' = f(x,y), \qquad y(x) \in \mathbb{R}^n, \quad f(x,y) \in \mathbb{R}^n.$$

Mögliche Normen in $\mathbb{R}^n$:

$$||y||_2 = (y_1^2 + ... + y_n^2)^{1/2} \qquad (euklidische\ Norm)$$

$$||y||_\infty = max_{1 \le i \le n}|y_i| \qquad (Maximum - Norm)$$

$$||y_1|| = |y_1| + |y_2| + ... + |y_n|$$

Satz 28 *Seien*

$$G = \{(x,y)|x_0 \le X,\ ||y - y_0|| \le b\}; \quad x - x_0 \le \frac{b}{A}; \quad ||f(x,y)|| \le A.$$

Dann gilt für das explizite Euler-Verfahren

$$y_{k+1} = y_k + (x_{k+1} - x_k)f(x_k, y_k), \quad k = 0, 1, ...$$

Folgendes:

1. $(x_i, y_i) \in G\ \forall i$
2. $||y_h(x) - y_0|| \le A \cdot |x - x_0|$
3. $||y_h - [y_0 + (x - x_0)f(x_0, y_0)]|| \le \epsilon|x - x_0|$, *falls* $||f(x,y) - f(x_0,y_0)|| \le \epsilon$

Form der Lipschitz-Bedingung für Vektorfall:

Skalarfall ($f \in C^1(G)$, Motivation: Differenzenquotient):

$$|f(x,y) - f(x,z)| \le L|y - z| \quad mit \quad L := \max_{(x,y)\in G} \frac{\partial f}{\partial y}.$$

Aus der Taylor-Reihe für $f(x,y) \in \mathbb{R}^n$

$$f(x,y) = f(x,z) + J(x,\xi)(y - z)$$

mit Jacobi-Matrix $J(x,\xi) \in \mathbb{R}^{n\times n}$, $J_{ij} = \frac{\partial f_i}{\partial y_j}$ erhalten wir

$$||f(x,y) - f(x,z)|| \le ||J(x,\xi)||\ ||y - z|| \le L||y - z||, \qquad L := \max_{(x,y)\in G} ||J(x,y)||.$$

Matrixnormen

Axiome:

1. $|||A||| \ge 0$, und $|||A||| = 0 \Leftrightarrow A = 0$

2. $|||cA||| = |c| \; |||A|||$
   $\forall c \in \mathbb{R}$

3. $|||A + B||| \leq |||A||| + |||B|||$ (Dreiecksungleichung)

4. $|||AB||| \leq |||A||| \; |||B|||$ (Sub-Multiplikativität)

Definition 29 *(Induzierte Norm) Induzierte Norm ist ein Matrix-Norm, die einer Vektronorm wie folgt abgeleitet ist:*

$$|||A||| := \max_{||x||=1} ||Ax|| = \sup_{||x||\neq 0} \frac{||Ax||}{||x||}.$$

Induzierte Matrixnormen sind stets submultiplikativ:

$$||Ax|| \leq |||A||| \; ||x||.$$

Bemerkung 30 *Alle Matrixnormen sind äquivalent als spezielle Vektornormen (über endlichdimensionalem Vektorraum).*

4 Weitere Einschrittverfahren: Runge-Kutta Methoden

Das AWP

$$y'(x) = f(x, y(x)), \quad y(x_0) = y_0, \quad y(x) \in \mathbb{R}^n, \quad f \in \mathbb{R}^n$$

fuhrt auf entsprechende Integralgleichung

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\mathrm{dt}.$$

Das exsplizite Euler'sche Polygonzugverfahren ist eine sehr einfache Möglichkeit,

$$\int_{x_0}^{x} f(t, y(t))\mathrm{dt} \approx (x - x_0) f(x_0, y_0)$$

zu approximieren (Rechtescksregel mit der Auswertung im linken Intervallende).

Implizites Euler-Verfahren:

$$\int_{x_0}^{x} f(t, y(t))\mathrm{dt} \approx (x - x_0) f(x, y(x))$$

Auswertung im rechten Intervallende:

$$y_{k+1} = y_k + h_k f(x_{k+1}, y_{k+1}).$$

In jedem Schritt ist eine nichtlineare Gleichung zu lösen.

Definition 31 *(ESV)*
Allgemeine Form von Einschrittverfahren (ESV):

$$y_{k+1} = y_k + h_k F(x_k, y_k, y_{k+1}, h_k)$$

mit der Verfahrensfunktion F.
Explizites ESV: F hängt nicht von y_{k+1} ab.
Implizites ESV: F hängt von y_{k+1} ab.

4.1 Grundbegriffe und Definitionen, Butcher-Schema

Sei h eine Diskretisierung des Intervalls $[x_0, X]$ mit dem Gitter

$$x_0 < x_1 < ... < x_n = X; \quad G_h = \{x_0, ..., x_{n(h)}\}$$

und den Schrittweiten

$$h_i = x_{i+1} - x_i; \quad i = 0, ..., n-1;$$

$$h = \max_{0 \le i \le n-1} h_i$$

Definition 32 *(Konvergenz)*

1. *Sei $x \in [x_0, X]$ fix und Gitterpunkt für alle Diskretisierungen h, d.h. $x = x_{m(h)}$. Das Verfahren 31 konvergiert in x für $h \to 0$ wenn*

$$|y(x) - \underbrace{y_{m(h)}}_{y_h(x_{m(h)})}| \overset{h\to 0}{\longrightarrow} 0.$$

2. *Das Verfahren 31 konvergiert auf $[x_0, X]$, wenn für die Folge der Diskretisierungen gilt:*

$$\max_{x_i \in G_h} |y(x) - y_h(x_i)| \overset{h\to 0}{\longrightarrow} 0$$

y definiert auf $[x_0, X]$
y_h definiert auf G_h ("Gitterfunktion")
("Gleichmäßige Konvergenz auf dem Gitter")

3. *das Verfahren 31 gat die Konvergenzordnung $p > 0$, wenn*

$$\max_{x_i \in G_h} |y(x) - y_h(x_i)| = O(h^p) \ für \ h \to 0.$$

Bemerkung 33 *Das grundsätzliche Problem bei dieser Definition: Konvergenz einer Folge von Gitterfunktionen $y_h(x_i)$ (auf jeweils unterschiedlichem Gitter G_h) gegen stetige Funktion.*

Fehler des Verfahrens ("globaler Diskretisierungsfehler")

$$z_m = \underbrace{y(x_m)}_{exakte\ Lösung} - \underbrace{y_m}_{=y_h(x_m)}$$

erfüllt (z.B. für explizites ESV):

$$z_{m+1} = y(x_{m+1}) - \overbrace{[y(x_m) + hF(x_m, y(x_m), h_m)]}^{1.\ Approx.-Schritt\ ausgehend\ von\ exaktem\ y(x_m)} +$$

$$+[y(x_m) + hF(x_m, y(x_m), h_m)] - \underbrace{[y(x_m) + hF(x_m, y_m, h_m)]}_{=y_{m+1}} =$$

$$= \underbrace{h_m \psi_m^{(1)}}_{"neuer"\ Diskretisierungsfehler} + \underbrace{(z_m + h_m \psi_m^{(2)})}_{Fehlerfortpflanzung}.$$

- $\psi_m^{(1)} := \frac{1}{h_m}\{y(x_{m+1}) - [y(x_m) + hF(x_m, y(x_m), h_m)]\}$ heißt *lokaler Diskretisierungsfehler* an Stelle x_m, auch "**Residuum**"
- $\psi_m^{(2)} := F(x_m, \underbrace{y(x_m)}_{y_m + z_m}, h_m) - F(x_m, y_m, h_m)$

- *Fehler-Evolution*:
$$\frac{z_{m+1} - z_m}{h_m} = \psi_m^{(1)} + \psi_m^{(2)}$$

Definition 34 *(Konsistenz, Konsistenzordnung)*
Das Verfahren 31 ist

- *konsistent mit der Gleichung* $y' = f(x, y)m$ *wenn*
$$\max_{x_m \in G_h} |\psi_m^{(1)}| \to 0 \quad f\ddot{u}r \quad h \to 0;$$

- *konsistent (mind.) mit der Ordnung p, wenn*
$$\max_{x_m \in G_h} |\psi_m^{(1)}| = O(h^p), \quad h \to 0.$$

Bemerkung 35

- *Bei ESV ist Konsistenz für Konvergenz ausreichend, und es gilt: Konsistenzordnung p Rightarrow Konvergenzordnung p*

- *Bei Mehrschrittverfahren muss das Verfahren auch stabil sein.*

Beispiel 36 *("explizites Euler-Verfahren")*
Welche Konsistenzordnung hat das explizite Euler-Verfahren?

$$\int_{x_m}^{x_m+1} f(x, y(x))dx \approx (x_{m+1} - x_m)f(x_m, y_m) \qquad (Rechecksregel)$$

$$\Rightarrow F = f(x_m, y_m)$$

$$y(x_{m+1} = y(x_m) + y'(x_m)h_m + \frac{1}{2}y''(\xi)h_m^2,$$

d.h.
$$\psi_m^{(1)} = \underbrace{-f(x_m, y(x_m))) + y'(x_m)}_{=0} + \frac{1}{2}y''(\xi)h_m = O(h_m)$$

(falls $|y''(\xi)| \leq C$*), d.h. das Verfahren ost erster Ordnung.*

Beispiel 37 *("ein impliziertes Euler-Verfahren")*

$$\int_{x_m}^{x_m+1} f(x, y(x))dx \approx \frac{1}{2}(x_{m+1} - x_m)[f(x_m, y_m) + f(x_{m+1}, y_{m+1})] \; (Trapez-Regel, \; "Crank-Nicolson")$$

$$\Rightarrow F = \frac{1}{2}[f(x_m, y_m) + f(x_{m+1}, y_{m+1})]$$

Das Verfahren ist implizit, d.h. y_{m+1} *wird aus der nichtlinearen Gleichung*

$$y_{m+1} - \frac{h_m}{2}f(x_{m+1}, y_{m+1}) = y_m + \frac{h_m}{2}f(x_m, y_m)$$

bestimmt (z.B. mit Newton-Verfahren).

$$\psi_m^{(1)} = -\frac{1}{2}[f(x_m, y_m) + f(x_{m+1}, y_{m+1})] + \frac{y(x_{m+1}) - y(x_m)}{h_m} =$$

$$= -\frac{1}{2}[y'(x_m) + \underbrace{y'(x_{m+1})}_{=y'(x_m)+y''(x_m)h_m+O(h_m^2)}] + y'(x_m) + \frac{1}{2}y''(x_m)h_m + O(h_m^2) =$$

$$= O(h_m^2) \Rightarrow \ Konsistenzordnung = 2$$

Beispiel 38 *("Runge-Kutta-Verfahren der Ordnung 2", "modifiziertes Euler-Verfahren")*

Idee:

$$\int_{x_m}^{x_m+1} f(x, y(x))dx \approx h_m \cdot f\left(x_m + \frac{1}{2}h_m, y\left(x_m + \frac{1}{2}h_m\right)\right) \quad (Mittelpunktregel)$$

$$y\left(x_m + \frac{1}{2}h_m\right) \approx y_{m+\frac{1}{2}} = y_m + \frac{1}{2}h_m f(x_m, y_m) \quad \leftarrow \quad expl.\ Euler - Schritt$$

$$\Rightarrow F = f\left(x_m + \frac{1}{2}h_m, y_m + \frac{1}{2}h_m f(x_m, y_m)\right)$$

Konsistenzordnung:

$$\psi_m^{(1)} = -f\left(x_m + \frac{1}{2}h_m, y_m + \frac{1}{2}h_m f(x_m, y_m)\right) + \underbrace{\frac{y(x_{m+1} - x_m)}{h_m}}_{\downarrow} =$$

$$= -f(x_m, y(x_m)) - \frac{1}{2}h_m f_x(x_m, y(x_m)) -$$

$$-\frac{1}{2}h_m f(x_m, y(x_m)) f_y(x_m, y(x_m)) \quad + \quad \overbrace{y'(x_m) + \frac{1}{2}h_m y''(x_m) + O(h_m^2)}^{\uparrow} =$$

$$= O(h_m^2),$$

weil $y''(x_m) = f_x(x_m, y(x_m)) + f_y(x_m, y(x_m))f(x_m, y(x_m))$.

$\Rightarrow$ *Konsistenzordnung 2, zwei Funktionsaufrufe, explizites Verfahren!*

Algorithmische Realisierung:

$$k_1 := f(x_m, y_m)$$

$$k_2 := f(x_m + \frac{1}{2}h_m, y_m + \frac{1}{2}h_m k_1)$$

$$y_{m+1} = y_m + h_m \cdot k_2$$

Definition 39 *(s-stufiges Runge-Kutta-Verfahren)*
Sei $s \in \mathbb{N}$,

$$A = \begin{pmatrix} 0 & ... & ... & ... & 0 \\ a_{21} & 0 & ... & ... & 0 \\ a_{31} & a_{32} & 0 & ... & 0 \\ ... & ... & ... & ... & ... \\ a_{s1} & ... & ... & a_{s,s_1} & 0 \end{pmatrix} \in \mathbb{R}^{s \times s}$$

eine reelle Matrix, und

$$c = (0, c_2, c_3, ..., c_s)^T \in \mathbb{R}^s,$$

$$b = (b_1, b_2, ..., b_s)^T \in \mathbb{R}^s$$

zwei reelle Vektoren. Dann heißt das Verfahren

$$\begin{aligned} k_1 &= f(x_m, y_m) \\ k_2 &= f(x_m + c_2 h_m, y_m + h_m a_{21} k_1) \\ &... \\ k_s &= f(x_m + c_s h_m, y_m + h_m(a_{s1} k_1 + ... + a_{s,s-1} k_{s-1})) \\ y_{m+1} &= y_m + h_m(b_1 k_1 + ... + b_s k_s) \end{aligned}$$

s-stufiges Runge-Kutta-Verfahren (s - Anzahl der Funktionsaufrufe).

Beispiel 40 *Explizites Euler-Verfahren:*
$s = 1$, $A = 0$, $c = 0$, $b = 1$.

Beispiel 41 *RK 2.Ordnung (modifiziertes Euler-Verfahren):*
$s = 2$, $A = \begin{pmatrix} 0 & 0 \\ 1/2 & 0 \end{pmatrix}$, $c = (0, 1/2)^T$, $b = (0, 1)^T$

$$F = f(x_m + \frac{1}{2}h_m, y_m + \frac{1}{2}h_m \underbrace{f(x_m, y_m)}_{k_1}) = k_2$$

Definition 42 ***Schreibweise:***

$$\left(\begin{array}{c|c} c & A \\ \hline & b^T \end{array}\right) \quad \leftarrow \quad \textit{"}\textbf{\textit{Butcher-Schema}}\textit{" (1964)}$$

Z.B. RK-Verfahren 2. Ordnung wird wie folgt aufgeschrieben:

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

Handelt es sich um ein explizites RK-Verfahren, so werden die Nullen im oberen Dreieck der Matrix oft gar nicht geschrieben.

4.2 Runge-Kutta Verfahren 2. Ordnung

Alle RK-Verfahren fur $s = 2$ haben die Gestalt

$$\begin{aligned} k_1 &= f(x_m, y_m) \\ k_2 &= f(x_m + c_2 h_m, y_m + h_m a_{21} k_1) \\ y_{m+1} &= y_m + h_m(b_1 k_1 + b_2 k_2), \end{aligned}$$

beinhaltet also 4 Parameter $a_{21},\ b_1,\ b_2,\ c_2$.

Das Ziel ist, die Parameter so zu wählen, daß möglichst hohe Konsistenzordnung erreicht wird:

$$\psi_1^{(m)} = \frac{y(x_{m+1}) - y(x_m)}{h_m} - F(...) =$$

$$= \frac{y(x_{m+1}) - y(x_m)}{h_m} - b_1 f(x_m, y(x_m)) - b_2 f(x_m + c_2 h_m, y(x_m) + h_m a_{21} f(x_m, y(x_m)))$$

Mit Taylor-Entwicklung

$$\frac{y(x_{m+1}) - y(x_m)}{h_m} = y' + \frac{h_m}{2} \overbrace{y''}^{=f_x + f f_y} + O(h^2)$$

erhalten wir

$$f(x_m, y(x_m)) - b_2 f(x_m + c_2 h_m, y(x_m) + h_m a_{21} f(x_m, y(x_m))) =$$

$$= f(x_m, y_m) + c_2 h_m f_x(x_m, y_m) + a_{21} h_m f(x_m, y_m) f_y(x_m, y_m) + O(h^2).$$

Also

$$\Rightarrow \psi_m^{(1)} = y' - (b_1 + b_2) f - h_m \left[(b_2 a_{21} - \frac{1}{2}) f f_y + (b_2 c_2 - \frac{1}{2}) f_x \right] + O(h^2)$$

- für $b_1 + b_2 = 1$ (Konsitenzbedingung) $\Rightarrow$ *mindestens 1. Ordnung*
- für $b_1 + b_2 = 1$, $b_2 a_{21} = \frac{1}{2}$, $b_2 c_2 = \frac{1}{2}$ $\Rightarrow$ 2.Konsistenzordnung
  $\rightarrow$ 4 Variablen, 3 Gleichungen
  $\Rightarrow$ *einparametrige Familie der RK-Verfahren der Ordnung 2*:
  Sei $b_2 = \sigma \Rightarrow b_1 = 1 - \sigma,\ a_{21} = \frac{1}{2\sigma}, c_2 = \frac{1}{2\sigma}$.
- **Verfahren**:

$$y_{m+1} = y_m + h_m F,$$

$$F = (1 - \sigma) f(x_m, y_m) + \sigma f \left(x_m + \frac{1}{2\sigma} h_m, y_m + \frac{1}{2\sigma}) h_m f(x_m, y_m) \right)$$

- Für $\sigma = 1 \Rightarrow$ modifiziertes Euler-Verfahren, Runge 1895 (Beispiel 38)

- Für $\sigma = \frac{1}{2} \Rightarrow$ "Verfahren von Heun", 1900:

$$\begin{aligned} k_1 &= f(x_m, y_m) \\ k_2 &= f(x_m + \overbrace{c_2}^{=1} h_m, y_m + h_m \overbrace{a_{21}}^{=1} k_1) \\ y_{m+1} &= y_m + \frac{1}{2} h_m \underbrace{(k_1 + k_2)}_{b_1 = b_2 = \frac{1}{2} 1}. \end{aligned}$$

  Das ist eine explizite Variante des impliziten Euler-Verfahrens (Bsp. 37, Trapez-Regel)

Satz 43

Es existiert kein 2-stufiges RK-Verfahren der Ordnung 3.

Frage: Wie hoch ist die mit einem s**-stufigen Verfahren** erreichbare **maximale Konsistenzordnung** p?

s	1	2	3	4	5	6	7	8	≥ 9
$n(s)$	1	3	6	11	21	42	91	207	
$p^*(s)$	1	2	3	4	4	5	6	6	$\leq s-2$

Wobei

- $n(s)$ die Anzahl der Bedingungsgleichungen für größtmögliches p (nach elimination der redundanten Gleichungen) und
- $p^*(s)$ die maximal mögliche Konsistenzordnung eines s-stufiges RK-Verfahrens sind.

Nachteil der RK-Methoden: höhere Konsistenzordnung nur durch mehr Funktionsauswertungen.

4.3 Runge-Kutta Verfahren der Ordnung 3 und 4

Das Butcher-Schema für $s = 3$ hat die Form:

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ c_2 & a_{21} & 0 & 0 \\ c_3 & a_{31} & a_{32} & 0 \\ \hline & b_1 & b_2 & b_3 \end{array} \Rightarrow \ 8 \ \textit{Parameter.}$$

RK-Verfahren (mit Notation $h = h_m$):

$$\begin{aligned} k_1 &= f(x_m, y_m) =: f \\ k_2 &= f(x_m + c_2 h, y_m + h a_{21} k_1) \\ k_3 &= f(x_m + c_3 h, y_m + h a_{31} k_1 + h a_{32} k_2). \end{aligned}$$

Durch Taylor-Entwicklung von f um (x_m, y_m):

$$k_2 = f + c_2 h f_x + h a_{21} f_y k_1 + \frac{h^2}{2}(c_2^2 f_{xx} + 2 f_{xy} k_1 c_2 a_{21} + a_{21}^2 f_{yy} k_1^2) + O(h^2).$$

Mit $k_1 = f$, $k_2 = f + c_2 h f_x + a_{21} h f_y x f + O(h^2)$ erhält man:

$$k_3 = f + h(c_3 f_x + (a_{31} + a_{32}) f f_y) + \frac{h^2}{2}[c_3^2 f_{xx} + 2c_3(a_{31} + a_{32}) f f_{xy} +$$
$$+ 2(a_{32} c_2 f_x f_y + a_{32} a_{21} f f_y^2) + (a_{31} + a_{32})^2 f^2 f_{yy}] + O(h^3).$$

Damit gilt:

- $$F = b_1 k_1 + b_2 k_2 + b_3 k_3 =$$
$$= \quad h^0 \cdot (b_1 + b_2 + b_3) f +$$
$$+ h^1 \cdot [(c_2 b_2 + c_3 b_3) f_x + (b_2 a_{21} + b_3(a_{31} + a_{32})) f f_y] +$$
$$+ h^2 \cdot \frac{1}{2}[(b_2 c_2^2 + b_3 c_3^2) f_{xx} + 2(b_2 c_2 a_{21} + b_3 c_3(a_{31} + a_{32})) f f_{xy} +$$
$$+ 2 b_3 c_2 a_{32} f_x f_y + 2 b_3 a_{32} a_{21} f f_y^2 + (b_2 a_{21}^2 + b_3(a_{31} + a_{32})^2) f^2 f_{yy}] +$$
$$+ O(h^3).$$

- $$\frac{y(x_{m+1}) - y(x_m)}{h_m} = y'(x_m) + \frac{h_m}{2} y''(x_m) + \frac{h_m^2}{6} y'''(x_m) + O(h_m^3) =$$
$$= f + \frac{h}{2}(f_x + f f_y) + \frac{h^2}{6}(f_{xx} + 2 f_{xy} f + f_{yy} f^2 + f_x f_y + f_y^2 f) + O(h_m^3).$$

- Bedingungen für $\psi_m^{(1)} = \frac{y(x_{m+1}) - y(x_m)}{h_m} - F \stackrel{!}{=} O(h^3)$ sind dann:

  1. $$b_1 + b_2 + b_3 = 1$$
  ist die Konsistenzbedingung, mindestens Ordnung 1.

  2. $$c_2 b_2 + c_3 b_3 = b_2 a_{21} + b_3(a_{31} + a_{32}) = \frac{1}{2}$$
  liefert mindestens die Ordnung 2.

  3. $$b_2 c_2^2 + b_3 c_3^2 = b_2 c_2 a_{21} + b_3 c_3(a_{31} + a_{32}) = b_2 a_{21}^2 + b_3(a_{31} + a_{32})^2) = \frac{1}{3}$$
  $$b_3 c_2 a_{32} = b_3 a_{32} a_{21} = \frac{1}{6}$$
  sind für die Konsistenzordnung 3 notwendig.

Insgesamt hat man 8 Gleichungen mit 8 Variablen. Die Gleichungen sind aber rednundant, so daß es nur 6 unabhängigen Gleichungen bleiben → Parameter sind nicht eindeutig bestimmt. Nach der Reduzierung

$$\begin{aligned} b_1 &= 1 - b_2 - b_3 \\ c_2 &= a_{21} \\ c_3 &= a_{31} + a_{32} \end{aligned}$$

hat man 3 Gleichungen mit 5 Variablen

$$\left\{ \begin{array}{lcl} c_2 b_2 + c_3 b_3 & = & \frac{1}{2} \\ c_2^2 b_2 + c_3^2 b_3 & = & \frac{1}{3} \\ c_2 b_3 a_{32} & = & \frac{1}{6} \end{array} \right.$$

Bemerkung 44

1. *$b_3 \neq 0$, $c_2 \neq 0$, $a_{32} \neq 0$, $a_{21} \neq 0$ notwendig.*
2. *für maximale Konsistenzordnung muss immer $c_i = \sum_{j=1}^{i-1} a_{ij}$*

$\longrightarrow$ R.-K.-Methode:

$$\begin{aligned} k_1 &= f(x_m, y_m) \\ k_2 &= f(x_m + c_2 h_m, y_m + \overbrace{c_2}^{=a_{21}} h_m k_1) \\ k_3 &= f(x_m + c_3 h_m, y_m + \underbrace{c_3}_{=a_{31}+a_{32}} h_m k_1 + \frac{h_m}{\underbrace{6 b_3 c_2}_{a_{32}=\frac{1}{6 b_3 c_2}}}(k_2 - k_1)) \\ y_{m+1} &= y_m + h_m(\underbrace{(1 - b_2 - b_3)}_{=b_1} k_1 + b_2 k_2 + b_3 k_3) \end{aligned}$$

wobei $C \cdot \begin{pmatrix} b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{3} \end{pmatrix}$ mit $C := \begin{pmatrix} c_2 & c_3 \\ c_2^2 & c_3^2 \end{pmatrix}$.

$\Rightarrow$ *Zweiparametrige Familie* von R.-K.-Metoden der Ordnung 3:

- z.B. c_2, c_3 fri aber $c_2 \neq 0$, $c_3 \neq 0$, $c_2 neq c_3$ (sonst C singular)

$$\Rightarrow b_2 = \frac{3c_2 - 2}{6c_2(c_3 - c_2)};\ b_3 = \frac{2 - 3c_2}{6c_3(c_3 - c_2)}$$

- für $c_3 = 0 \Rightarrow c_2 = \frac{2}{3},\ b_2 = \frac{3}{4},\ b_3$ frei
  $\Rightarrow$ einparametrige Familie.
- für $c_2 = c_3 = \frac{2}{3} \Rightarrow b_2 + b_3 = \frac{3}{4}$ → eine andere einparametrige Familie der R.-K.-Methoden.

Beispiel 45

1.

0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	0	0
1	-1	2	0
	$\frac{1}{6}$	$\frac{4}{6}$	$\frac{1}{6}$

Simpson-Formel (=Newton-Côtes Formel für 3 äquidistante Knoten):

$$\int_{x_m}^{x_{m+1}} f(x,y)dx \approx h_m \cdot \left[\frac{1}{6} f(x_m, y(x_m)) + \frac{4}{6} f(x_m + \frac{1}{2}h_m, y(x_m + \frac{1}{2}h_m)) + \frac{1}{6} f(\underbrace{x_{m+1}}_{x_m + 1 \cdot h_m}, y(x_{m+1})) \right]$$

2.

0	0	0	0
$\frac{1}{3}$	$\frac{1}{3}$	0	0
$\frac{2}{3}$	0	$\frac{2}{3}$	0
	$\frac{1}{4}$	0	$\frac{3}{4}$

Methoden 4. Ordnung ($s = p = 4$)

Klassische R-K.-Formeln (Kutta, 1901):

0				
$\frac{1}{2}$	$\frac{1}{2}$			
$\frac{1}{2}$	0	$\frac{1}{2}$		
1	0	0	1	
	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

$\frac{3}{8}$-Regel (Kutta, 1901):

0				
$\frac{1}{3}$	$\frac{1}{3}$			
$\frac{2}{3}$	$-\frac{1}{3}$	1		
1	1	-1	1	
	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

→ Newton'sche $\frac{3}{8}$-Regel (=Newton-Côtes Formel für 4 äquidistante Knoten).

4.4 Konvergenzresultate für die R.-K.-Methoden

Eine R.-K.-Methode hat die Form ($h_m = h = const$)

$$\begin{cases} \frac{y_{m+1}-y_m}{h} & = & \sum_{i=1}^{s} b_i k_i = F(y_m)) \\ k_1 & = & f(x_m, y_m), \quad \text{d.h. } k_i = k_i(y_m) \\ k_2 & = & f(x_m + c_i h, y_m + h\sum_{j=1}^{i-1} a_{ij} k_j),\ i = 2, ..., s \end{cases}$$

- Der Fehler $z_m = y(x_m) - y_m$ erfüllt die Gleichung $\frac{z_{m+1}-z_m}{h} = \psi_m(1) + \psi_m^{(2)},\ m = 0, 1, ...$ mit $\psi_m(1) = \frac{y(x_{m+1})-y(x_m)}{h} - \sum_{i=1}^{s} b_i k_i(y(x_m))$ $\psi_m(2) = F(y(x_m)) - F(y_m) = \sum_{i=1}^{s} b_i[k_i(y(x_m)) - k_i(y_m)]$
- zusätzlich: $z_0 = y(x_0) - y_0 = 0$.
- $x_m = x_0 + m \cdot h \leq X = const.$
- Die Funktion $f(x, y)$ erfülle auf G eine L-Bedingung bgzl. y.

$\Rightarrow$ fur $i \geq 2$:
$r_i := \|k_i(y(x_m))) - k_i(y_m)\| =$

$$= \left\| f\left(x_m + c_i h, y(x_m) + h\sum_{j=1}^{i-1} a_{ij} k_j(y(x_m))\right) - f\left(x_m + c_i h, y_m + h\sum_{j=1}^{i-1} a_{ij} k_j(y_m)\right)\right\| \leq$$

$$\leq L\left[\|y(x_m) - y_m\| + h\sum_{j=1}^{i-1} |a_{ij}| \|k_j(y(x_m)) - k_j(y_m)\|\right] \leq$$

$$\leq L\|z_m\| + Lha\sum_{j=1} i - 1 r_j, \text{ mit } a = \max_{\substack{2\leq i\leq s \\ 1\leq j\leq i-1}} |a_{ij}|$$

$$\overbrace{\|z_m\|}^{gesucht!}$$

und damit $r_{i+1} \leq Lha\sum_{j=1}^{i} r_j + L\ \overbrace{\|z_m\|}^{gesucht!}$ $\rightarrow$ rekursive Abschätzung!

Satz 46 *Die Funktion $f(x, y)$ sei auf G stetig und erfülle eine L-Bedingung bzgl. y. Für den Fehler einer s-stufigen R.-K.-Methode gilt dann:*

$$\|z_m\| = \|y(x_m) - y_m\| \leq (X - x_0)e\alpha(X - x_0) \cdot \max_{0\leq j\leq m} \| \overbrace{\psi_m^{(1)}}^{lok.\ Diskret.-Fehler} \|,$$

wobei

$$\alpha := sbLe^{Lah_0(s-1)},\ h \leq h_0$$

$$a := \max_{\substack{2\leq i\leq s \\ 1\leq j\leq i-1}} |a_{ij}|,$$

$$b := \max_{1\leq i\leq s} |b_i|.$$

4.5 Schrittweitensteuerung, eingebettete R.-K.-Methoden

Ziel: Begrenzung des lokalen Diskretisierungsfehlers durch *Wahl einer geeigneten Schrittweite* h_m^* zur Verwendung *für nächsten Schritt*: $h_{m+1} = h_m^*$.

Bedingung:

$$\underbrace{|\psi_m(m)|h_m}_{y(x_{m+1})-y_{m+1}\ bei\ 'richtiger'\ AB\ an\ x_m} \leq TOL\ \forall m$$

$\psi_m(1)$ unbekannt

TOL vom Benutzer gewählte Genauigkeit, 'TOLeranz'

$\Rightarrow$ heuristische Fehlerschätzer durch Vergleich von 2 verschieden genauen Verfahren: y_m und $\hat{y}_m$ (wesentlich bessere Näherung, z.B. Ordnung h^{p+1} statt h^p), also: $|\hat{\psi}_m^{(1)}| \ll |\psi_m^{(1)}|$.

- *Fehlerschätzer* $\epsilon_m := |\hat{y}_m - y_m| \approx |y(x_m) - y_m|$
  nur für "zweitbestes" Verfahren (für den lokalen Fehler)
- Bedingung:
  (Verfahren y_m ist $p-ter$ Ordnung)

$$\epsilon \approx Ch_m p + 1 \ll TOL \quad \Rightarrow \quad C \approx \frac{\epsilon_m}{h_m p + 1}$$

- *akzeptable Schrittweite* h_m^*:
  $C(h_m^*)p + 1 \approx \rho \cdot TOL$, wobei ρ "sicherheitsfaktor" ist (z.B. $\rho = \frac{1}{4}$)
  $\Rightarrow (h_m^*)^{p+1} \approx \rho\frac{TOL}{C} \approx \frac{\rho TOL}{\epsilon_m} \cdot \underbrace{h_m^{p+1}}_{verwendete\ 'alte'\ Schrittweite}$
  $\Rightarrow$ *neue Schrittweite* (aus Fehlerschätzung):

$$h_{m+1} = h_m^* = \sqrt[p+1]{\frac{\rho TOL}{\epsilon_m}} h_m$$

- meist wird mit $\hat{y}_m$ weitergerechnet (genauer!)

Eingebettete Runge-Kutta-Verfahren

Ziel: effiziente Berechnung von $y_m; \hat{y}_m$ zur Fehlerschätzung

$\hat{y}_m$ - RKV der Ordnung $p+1$ $\quad \begin{array}{c|c} \hat{c} & \hat{A} \\ \hline & \hat{b}^T \end{array}$

y_m - RKV der Ordnung p $\quad \begin{array}{c|c} c & A \\ \hline & b^T \end{array}$

Typisch sind 5(4), 8(7) (Matlab-Funktionen ode45, ode87).

- *Problem*: zusätzliche Berechnung von $\hat{y}_m$ erhöht Aufwand stark.
  $\Rightarrow$ wähle $\hat{c} = c,\ \hat{A} = A, \hat{b} \neq b$
  $\Rightarrow$ min. Verfahrensstufe $s^* = p + 1$ kann nicht erreicht werden.
  Beispiel: RK4(3) (klassisches RKV ($p = s = 4$))

0					
$\frac{1}{2}$	$\frac{1}{2}$				
$\frac{1}{2}$	0	$\frac{1}{2}$			
1	0	0	1		
	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0
	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{1}{3}$

  Gleiche Koeffizienten machen $k_5^m = k_1^{m+1} \Rightarrow$ Aufwandsreduktion ("Fehlberg-Trick")
  b^T - Ergänzung zu $s = 5,\ p = 3$
  $k_5^m = f\left(x_{m+1}, y_m + h\left(\frac{1}{6}k_1^m + ... + \frac{1}{6}k_4^m\right)\right)$
  $y_{m+1} = y_m + h\left(\frac{1}{6}k_1^m + ... + \frac{1}{6}k_4^m\right)$
  $k_1^{m+1} = f(x_{m+1}, y_{m+1}) = k_5^m$
  $\Rightarrow$ *vertretbarrer Aufwand* für Fehlerschätzung

4.6 Implizite R.-K.-Methoden (ein Ausblick)

Die expliziten R.K.-Formeln sind ein Speziallfall von Formeln

$$k_i = f(x_m + c_i h_m, y_m + h_m \overbrace{\sum_{j=1}^{s}}^{bisher\ j \leq i-1} a_{ij} k_j); \quad i = 1, ..., s$$

$$y_{m+1} = y_m + h_m \sum_{i=1}^{s} b_i k_i; \qquad m = 0, 1, ...\ ,$$

die als *implizite R.-K.-Verfahren* bezeichnet werden.

Butcher-Schema:

c	A
	b^T

$c, b \in \mathbb{R}^s$,
$A \in \mathbb{R}s \times s$ evtl. voll besetzt.

Problem: Lösung eines nichtlinearen Glecihungssystems für $(k_1, ..., k_s) = K \in \mathbb{R}n \times s$ für alle $m = 0, 1, ...$

Vorteile:

1. mehr Parameter zur Verfügung $\Rightarrow$ höhere Konsistenzordnung (bis $p = 2s$ (!!!) für die *Gauß-Form* (wird nicht im Rahmen dieser Vorlesung behandelt))

2. günstige Stabilitätseigenschaften

2 Typen: Gauß-Form (bessere Konsistenzordnung) und Radau-Form (bessere Stabilitätseigenschaften als Gauß-Form).

5 Randwertprobleme

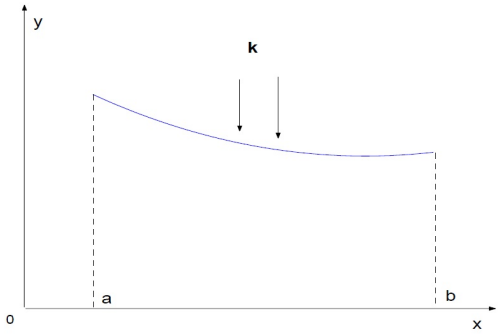
5.1 Aufgabestellung, Beispiele, Existenzresultate

Anwendungen

Beispiel 47 *Durchbiegung einer belasteten Saite:*

$$-(\mu(x)y')' = k(x),\ x \in (a,b)$$

$\mu(x) > 0$ *Spannung*
$k(x)$ *Kraftdichte*
$Gesamtkraft = \int_a^b k(s)ds$



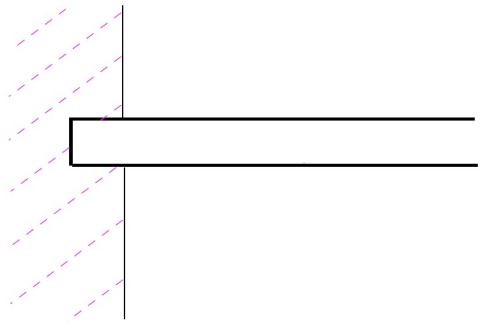
Randbedingungen:
$y(a) = \alpha,\ y(b) = \beta \rightarrow$ *RB 1. Art bzw. Dirichlet-RB*
$y'(a) = \alpha,\ y'(b) = \beta \rightarrow$ *RB 2. Art bzw. Neumann-RB*
$\left.\begin{array}{l} c_1y(a) + c_2y'(a) = \alpha \\ c_3y(b) + c_4y'(b) = \beta \end{array}\right\} \rightarrow$ *RB 3. Art bzw. Robin-RB*

/

Beispiel 48 *Durchbiegung eines elastischen Stabes:*

$$(\mu(x)y'')'' = k(x),\ x \in (a,b)$$

$\mu(x) > 0$ *Biegesteifigkeit*



RB:

- *frei drehbare Lagerung, z.B.:* $= 0y(a) = y(b) = y''(a) = y''(b)$
- *fest eingespannt, z.B.:* $0 = y(a) = y'(a)$
- *skalare Gleichung höherer Ordnung* $\rightarrow$ *auf System* 1. *Ordnung transformierbar*
- *allgemeine nichtlineare Form:*

$$\begin{cases} y' = f(x,y),\ x \in (a,b);\ y, f \in \mathbb{R}^n \\ r(y(a), y(b)) = 0,\ r \in \mathbb{R}^n \end{cases}$$

- *in der Praxis oft lineare RBen:*
  $r(y(a), y(b)) = Ay(a) + By(b) = C;\ A, B \in \mathbb{R}^{n\times n};\ C \in \mathbb{R}^n$
  oder separierte RBen:
  $Ay(a) = \underbrace{C_a}_{\in\mathbb{R}^r},\ By(b) = \underbrace{C_b}_{\in\mathbb{R}^s};\ A \in \mathbb{R}^{r\times n};\ B \in \mathbb{R}^{s\times n},\ r+s=n$

Beispiel 49 $y''(x) + y(x) = 0$
$\rightarrow$ *allgemeine Lösung:* $y(x) = c_1 \sin x + c_2 \cos x;\quad c_1, c_2 \in \mathbb{R}$

1. $y(0) = 0,\ y(\pi/2) = 1$
   $\rightarrow y(x) = \sin x$ *eindeutige Lösung*
2. $y(0) = y(\pi) = 0$
   $\rightarrow y(x) = c_1 \sin x,\ c_1 \in \mathbb{R}$ *beliebig viele Lösungen*
3. $y(0) = 0,\ y(\pi) = 1 \Rightarrow$ *keine Lösung!*

Rightarrow Existenz-eindeutigkeitstheorie ganz anders als bei AWP.

Beispiel 50 *Skalares lineares RWP 2. Ordnung*

$$\begin{cases} -y'' + p(x)y' + q(x)y = g(x),\ x \in (a,b) \\ y(a) = \alpha, y(b) = \beta \end{cases}$$

Bemerkung:

$$w(x) := y(x) - \alpha\frac{x-b}{a-b} + \beta\frac{x-a}{b-a}$$

transformiert das Problem auf RWP für w *mit homogenen Randbedingungen, d.h.* $w(a) = w(b) = 0$.

Satz 51 *(Fredholm-Alternative)*

1. *Das RWP aus Beispiel (50) ist genau dann für alle* $g \in C[a,b],\ (\alpha,\beta)^T \in \mathbb{R}^2$ *eindeutig lösbar, wenn das zugehörige homogene Problem (d.h.* $g, \alpha, \beta = 0$*) nur die triviale Lösung hat ("Fredholm-Alternative")*

2. Für $q(x) \geq 0$ hat das homogene Problem (50) nur die triviale losung.

Idee für weitere Existenzresultate

Betrachten wir das RWP

$$\begin{cases} y' = f(x,y),\ x \in (a,b) \\ r(y(a), y(b)) = 0 \end{cases} \tag{3}$$

Umformulierung als AWP:

$$\begin{cases} y' = f(x,y),\ x \in (a,b) \\ y(a) = s \in \mathbb{R}^n \end{cases} \Rightarrow \textit{ Lösung } y = y(x;s) \tag{4}$$

$y(x;s)$ löst auch (3), falls die Funktion

$$F(s) := r(s, y(b;s)) : \mathbb{R}^n \to \mathbb{R}^n$$

eine Nullstelle hat. Das Gilt z.B. falls

$$\Phi(s) := s - QF(s) : \mathbb{R}^n \to \mathbb{R}^n$$

eine Kontraktion ist, mit geeigneter regulärer Matrix $Q \in \mathbb{R}^{n \times n}$.

Der Fixpunkt s^* von Φ ist dann die Nullstelle von $F(s) \Rightarrow\ y(x;s^*)$ löst (3).

Eigenwertprobleme:

$$\begin{cases} y' = f(x,y,\lambda),\ x \in (a,b) \\ r(y(a), y(b), \lambda) = 0 \end{cases} \Rightarrow \textit{ Lösung } y = y(x;s) \tag{5}$$

- (5) ist typischerweise überbestimmt $\Rightarrow \exists$ keine oder nur triviale Lösung.
- Gesucht: $\lambda \in \mathbb{C}$ (=*Eigenwerte*), sodaß (5) eine nichttriviale Lösung $y(x)$ (=*eigenfunktion*) hat.

Beispiel 52 $\quad y' + \lambda y = 0, \lambda \in \mathbb{R},\ x \in (0,1)$, $y(0) = y(1) = 0$

1. $\lambda < 0$*:*

   Allgem. Lösung: $y(x) = c_1 e^{\sqrt{-\lambda}x} + c_2 e^{-\sqrt{-\lambda}x} \Rightarrow y(x) \equiv 0$

2. $\lambda = 0$*:*

   $y(x) = c_1 + c_2 x \Rightarrow y(x) \equiv 0$

3. $\lambda > 0$*:*

   $y(x) = c_1 \sin(\sqrt{\lambda}x) + c_2 \cos(\sqrt{\lambda}x)$

   $$\Rightarrow \begin{cases} c_2 = 0 \\ c_1 \sin\sqrt{\lambda} = 0 \end{cases}$$

$\Rightarrow \sqrt{\lambda} = \pi k,\ k \in \mathbb{N}$

$\Rightarrow$ *Eigenwerte:* $l_k = \pi^2 k^2$

Eigenfunktionen: $y_k(x) = c \sin(\pi k x)$

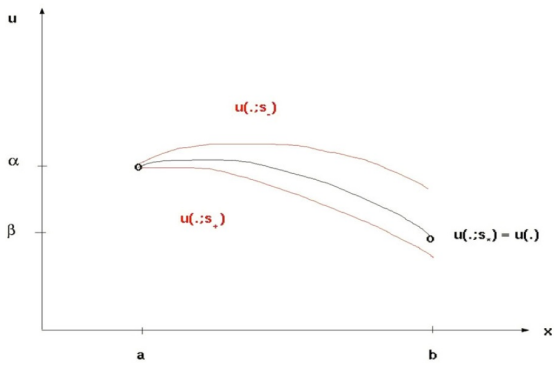
5.2 Einfach- und Mehrfachschießverfahren

5.2.1 Das Einfachschießverfahren ('Wasserschlauchmethode')

Gegeben sei das folgende RWP:

$$\begin{cases} u'' &= f(x, u, u'),\ a < x < b, \\ u(a) &= \alpha \\ u(b) &= \beta. \end{cases}$$

Annahme: $\exists!$ Lösung $u(x)$.



Das AWP

$$\begin{cases} u'' &= f(x, u, u'),\ a < x, \\ u(a) &= \alpha \\ u'(a) &= S. \end{cases}$$

besitzt dann i.A. eine eindeutige Lösung für alle s: $u(x) \equiv u(x; s)$.

Idee: Um das RWP zu lösen, müssen wir s so bestimmen, daß die 2. Randbedingung erfüllt ist, d.h. $u(b; s) = \beta$.

Man sucht damit eine Nullstelle s^* der Funktion

$$F(s) = u(b; s) - \beta.$$

Methoden:

1. Kennt man, z.B., Werte s_+ und s_- mit $F(s_+) > 0,\ F(s_-) < 0$, so kann man s^* mit *Bisektionsverfahren* berechnen (s. Skript 'Angewandte Mathematik').

2. *Newton-Tangentenverfahren* (s. Skript 'Angewandte Mathematik'): $u(b,s)$, und damit auch $F(s)$ sind stetig dierenzierbare Funktionen von s, also:

   $s_{i+1} = s_i - \frac{F(s_i)}{F'(s_i)}, \quad i = 0, 1, ...; \quad s_0$ gegeben (passend gewählt, so daß die Konvergenzbedingung erfüllt ist).

   $F(s_i)$ kann man durch lösung des AWPs bestimmen: Den Wert der *Ableitung* $F'(s_i)$

$$F'(s_i) = \frac{\partial}{\partial s} u(b; s)$$

   erhält man wie folgt:

   Die Funktion $v(x) = \frac{\partial}{\partial s} u(x; s)$ erfüllt die spezielle lineare Differentialgleichung

$$v'' = f_u(x, u, u')\frac{\partial u}{\partial s} + f_{u'}(x, u, u')\frac{\partial u'}{\partial s} =$$

$$= f_u(x, u, u')v + f_{u'}(x, u, u')v'$$

   mit AB:

$$v(a) = 0, \; v'(a) = 1.$$

   Beweis:

$$v'' = u_{sxx} = \frac{d}{ds} f\left(x, u(x; s), \frac{\partial u}{\partial x}(x; s)\right)$$

$$u(a; s) \equiv \alpha \to v(a) = 0$$

$$\frac{\partial u}{\partial x}(a; s) = s \to v'(a) = 1$$

   q.e.d.

   Das Problem wird damit oft sehr kompliziert. Man ersetzt daher oft die Ableitung $F'(s_i)$ durch einen Differenzenquotienten (s. Anhang 6.5), z.B.:

$$F'(s_i) \approx \Delta F(s_i) = \frac{F(s_i + \Delta s) - F(s_i)}{\Delta s}.$$

   Probleme:

   (a) $F(s_i + \Delta s)$ ist analog zu berechnen.

   (b) Δs zu groß $\Rightarrow$ schlechte Konvergenz des (naherungsweisen) Newton-Verfahrens.

   (c) Δs zu klein $\to F(s_i + \Delta s) \approx F(s_i) \Rightarrow$ Auslöschung
   $\Rightarrow F(s_i), \; F(s_i + \Delta s)$ müssen sehr genau berechnet werden.

(d) lokale quadratische Konvergenz geht i.A. verloren.

Bemerkung 53 *Sonderfall lineare RWPe:*

$$F'(s_i) = \Delta F(s_i) \quad \forall s$$

$\rightarrow F(s) = 0$ *in einem Schritt lösbar (lineares GS).*

Probleme des Schießverfahrens:

1. Oft instabil bezgl. 'Daten', z.B.
   (a) Max. Existenzintervall von $u(x; s)$ hängt i.A. von s ab $\rightarrow u(b; s)$ evtl. nicht definiert.
   (b) AWP numerisch instabil, obwohl RWP *gut konditioniert* ($\rightarrow$ mit Differenzenferfahren (s. Kapitel 5.3) stabil lösbar.

Beispiel 54

$$\left\{ \begin{array}{lcl} y_1'' & = & 110y_1 + y_1' \\ y_1(0) & = & y_1(10) = 1 \end{array} \right. \tag{6}$$

$$\rightarrow \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \underbrace{\begin{pmatrix} 0 & 1 \\ 110 & 1 \end{pmatrix}}_{=A} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

Allgemeine Lösung:

$$y(x) = c_1 e^{-10x} \begin{pmatrix} 1 \\ -10 \end{pmatrix} + c_2 e^{11x} \begin{pmatrix} 1 \\ 11 \end{pmatrix}, \quad c_1, c_2 \in \mathbb{R}$$

- *Das System ist nicht stabil*
  Randbedingung in der Matrixform:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} y(0) + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} y(1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

- *eindeutige Lösung des RWPs:*

$$y(x) = \underbrace{\frac{e^{110} - 1}{e^{110} - e^{-100}}}_{\approx 1} e^{-10x} \begin{pmatrix} 1 \\ -10 \end{pmatrix} + \underbrace{\frac{1 - e^{-100}}{e^{110} - e^{-100}}}_{\approx e^{-110}} e^{11x} \begin{pmatrix} 1 \\ 11 \end{pmatrix}$$

- *Anfangsbedingung:*

$$y(0) = s = \begin{pmatrix} 1 \\ -10 + \frac{21(1 - e^{-100})}{e^{110} - e^{-100}} \end{pmatrix} \approx \begin{pmatrix} 1 \\ -10 \end{pmatrix}$$

- *Bei* $10-$*stelliger Genauigkeit:*
  z.B.: $\tilde{s} = \begin{pmatrix} 1 \\ -10 + 10^{-9} \end{pmatrix}$
  Rightarrow Lösung $y_1(10; \tilde{s}) \approx \frac{10^{-9}}{21} e^{110} \approx 2.8 \cdot 10^{37}$
- *Sensitivität aus dem Satz 17 klar (Abhängigkeit von ABen):*

$$\begin{aligned} y' = f(x,y),\ y(x_0) = y_0 \\ \tilde{y}' = f(x,\tilde{y}),\ \tilde{y}(x_0) = \tilde{y}_0 \\ \Rightarrow |y(x) - \tilde{y}(x)| \leq e^{L(x-x_0)} |y_0 - \tilde{y}_0|. \end{aligned}$$

Hier mit $\lambda(A)$ *Eigenwert der Matrix A:*

$$\begin{aligned} L = 11 = |\lambda_{max}(A)| \\ \Rightarrow ||y(x;s) - y(x;\tilde{s})|| = O(e^{11x}) ||s - \tilde{s}|| \end{aligned}$$

Verbesserte Idee: Verkleinerung des Abstands $|x - a|$, also Verkleinerung des Intervalls $\rightarrow$ Motivation für das nächste Kapitel über das Mehrfachschießverfahren.

5.2.2 Das Mehrfachschießverfahren ('Multiple Shooting Method', MSV)

Gegeben sei das RWP:

$$y' = f(x,y),\ r(y(a), y(b)) = 0,\ y \in \mathbb{R}.$$

Idee: Bei MSV werden die Werte

$$s_k = y(x_k),\ k = 0, .., m$$

an mehreren Stellen

$$a = x_0 < x_1 < ... < x_m = b$$

gleichzeitig iterativ berechnet. Sei dazu

$$y(x; x_k, s_k)$$

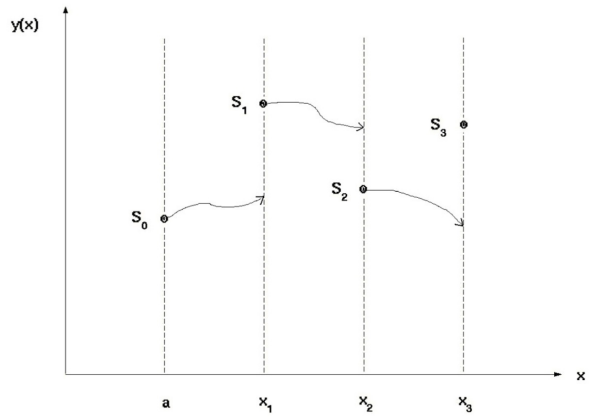
die Lösung des AWPs

$$y' = f(x,y),\ x > x_k,\ y(x_k) = s_k \in \mathbb{R}^n.$$

Aufgabe:
Bestimme die Vektoren s_k so, daß die stückweise zusammengesetzte Funktion

$$\begin{cases} y(x) &= y(x; x_k, s_k), \quad x \in [x_k, x_{k+1}) \\ y(b) &= s_n \end{cases}$$

stetig ist und die **Randbedingung** $r(s_0, s_m) = 0$ erfullt.



- Es ergeben sich damit $n \cdot (m+1)$ Bedingungen:

$$\begin{cases} y(x_{k+1}; x_k, s_k) &= s_{k+1}, \quad k = 0, ..., m-1 \\ r(s_0, s_m) &= 0 \end{cases}$$

- Das Gleichungssystem für die $n \cdot (m+1)$ Unbekannten $s_0, ..., s_m$ lautet mit $S = (s_0, ..., s_m)^T \in \mathbb{R}^{(m+1)}$:

$$F(S) := \begin{pmatrix} y(x_1; x_0, s_0) - s_1 \\ \vdots \qquad \vdots \qquad \vdots \\ y(x_m; x_{m-1}, s_{m-1} - s_m) \\ r(s_0, s_m) \end{pmatrix} = 0 \in \mathbb{R}^{n \cdot (m+1)}$$

  und kann mit Hilfe des Newton-Verfahrens

$$S_{i+1} = S_i - [DF(S_i)]^{-1} F(S_i), \ i = 0, 1, ...$$

  iterativ gelöst werden.

- *Berechnung von* $F(S_i)$: m AWPe zu lösen.

- Die *Matrix* $DF(S) = \left(\frac{\partial}{\partial s_k} F_i(S)\right)_{i,k=0,...,m}$ hat die folgende Block-Gestalt:

$$DF(S) = \begin{pmatrix} G_0 & -I & 0 & ... & ... & 0 \\ 0 & G_1 & -I & 0 & ... & 0 \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & G_{m-1} & -I \\ A & 0 & ... & ... & 0 & B \end{pmatrix}$$

mit einzelnen Matrix-Blöcken:

$$I = diag(1, ..., 1) \in \mathbb{R}^{n\times n} \ (n \times n - Einheitsmatrix)$$
$$G_k = D_{s_k} F_k(S), \quad k = 0, ..., m-1; \qquad G_k \in \mathbb{R}^{n\times n}$$
$$A = D_{s_0} r(s_0, s_m) \in \mathbb{R}^{n\times n}$$
$$B = D_{s_m} r(s_0, s_m) \in \mathbb{R}^{n\times n}$$

$\Rightarrow$ Newton-Verfahren lautet mit $\Delta S = (\Delta s_0, ..., \Delta s_m)^T = S_{i+1} - S_i$:

$$DF(S)\Delta S = -F,$$

oder, mit der Kenntnissnahme der Struktur der Matrix $DF(S)$:

$$\begin{cases} G_0\Delta s_0 - \Delta s_1 = -F_0 \\ G_1\Delta s_1 - \Delta s_2 = -F_1 \\ \vdots \qquad \vdots \qquad \vdots \\ G_{m-1}\Delta s_{m-1} - \Delta s_m = -F_{m-1} \\ A\Delta s_0 + B\Delta s_m = -F_m \end{cases}$$

$\Rightarrow \Delta s_1 = G_0\Delta s_0 + F_0$, u.s.w.

$$\Delta s_m = \underbrace{G_{m-1}G_{m-2}...G_1G_0}_{=:G\in\mathbb{R}^{n\times n}}\Delta s_0 + \underbrace{\sum_{j=0}^{m-1}\left(\prod_{l=j+1}^{m-1} G_l\right)F_j}_{=:w\in\mathbb{R}^n} = G\Delta s_0 + w$$

$\Rightarrow$ lineares GS für Δs_0 :

$$(A + BG)\Delta s_0 = -F_m - Bw$$

$\rightarrow \Delta s_i$ werden aus Δs_0 sukzessive berechnet.

- *Praktische Aspekte:*
  - *Modifiziertes Newton-Verfahren* $\rightarrow$ nur Approximation der Matrizen $DF(S_i)$
  - *Wahl der Zwischenpunkte x_k:*
    Sei $z(x)$ eine *Starttrajektorie*
    * $x_0 := a$
    * Sei $x_i < b$ gewählt, so integriere das AWP

    $$\begin{cases} z_i' & = & f(x, z_i), \ x > x_i \\ z_i(x_i) & = & z(x_i) \end{cases}$$

    bis zu $x = \xi$, wo $||z_i(\xi)|| \geq \lambda||z(\xi)||$ (z.B. $z = 2$);
    setze $x_{i+1} := \xi$.
  - Falls $f(x, y)$ nur stückweise stetig auf $[a, b]$ bzw. stückweise stetig differenzierbar auf $[a, b]$ ist:
    $\rightarrow$ Unstetigkeiten als Zwischenpunkte wählen, sonst Konvergenzprobleme.

5.3 Differenzenverfahren

RWP:

$$\begin{cases} -y'' + q(x)y = g(x), \ x \in (a,b) \\ y(a) = \alpha, y(b) = \beta \end{cases}$$

- Seien $p, q \in C[a,b], \ q \geq 0 \Rightarrow \exists!$ Lösung $y \in C^2[a,b]$ (laut Fredholm-Alternative, Satz (51))
- Diskretisierung:

  Idee: Approximation der Ableitungen in der Differentialgleichung (s. Anhang)

  Gitterpunkte: $x_j = a + jh, \ j = 0, ..., n; \ h = \frac{b-a}{n}$
- *Approximation der 1. Ableitung:* (für $y \in C^2$ bzw. C^3):

  Vorwärtsdifferenz:

$$(D^+y)(x) := \frac{y(x+h) - y(x)}{h} = y'(x) + \overbrace{\frac{h}{2}}^{Approx.Ordn\ 1} y''(x + \underbrace{\theta_1}_{0 \leq \theta_1 \leq 1} h)$$

  Rückwärtsdifferenz:

$$(D^-y)(x) := \frac{y(x) - y(x-h)}{h} = y'(x) - \overbrace{\frac{h}{2}}^{Approx.Ordn\ 1} y''(x - \underbrace{\theta_2}_{0 \leq \theta_2 \leq 1} h)$$

  Symmetrische Differenz:

$$(D^0y)(x) := \frac{y(x+h) - y(x-h)}{2h} = y'(x) + \frac{h^2}{6} y'''(x + \underbrace{\theta}_{-1 \leq \theta \leq 1} h) \quad (Ordnung\ 2)$$

- *2. Ableitung*: (für $y \in C^4$):

  Zentraler Differenzenquotient 2. Ordnung:

$$(D^+D^-y)(x) := \frac{y(x+h) - 2y(x) + y(x-h)}{h^2} = y''(x) + \frac{h^2}{12} y''''(x + \underbrace{\theta}_{-1 \leq \theta \leq 1} h)$$

- *Diskretizierung des RWPs:*

  $y_j \approx y(x_j)$

  $q_j = q(x_j)$

  $g_j = g(x_j)$

$$\left\{ \begin{array}{l} y_0 = \alpha \\ -\frac{y_{j+1}-2y_j+y_{j-1}}{h^2} + q_j y_j = g_j; \quad j = 1, ..., n-1 \\ y_n = \beta \end{array} \right\} \leftarrow n{+}1\ Gl.\ für\ y_0, ..., y_n$$

bzw.

$$\left\{ \begin{array}{l} (2 + h^2 q_1) y_1 - y_2 = h^2 g_1 + \alpha \\ -y_{j-1} + (2 + h^2 q_j) y_j - y_{j+1} = h^2 g_j; \quad j = 2, ..., n-2 \\ -y_{n-2} + (2 + h^2 q_{n-1}) y_{n-1} = h^2 g_{n-1} + \beta \end{array} \right\} \leftarrow n{-}1\ Gl.\ für\ y_1, ..., y_{n-1}$$

Matix-Form:

$$\underbrace{\begin{pmatrix} 2+h^2q_1 & -1 & 0 & ... & 0 \\ -1 & 2+h^2q_2 & -1 & ... & 0 \\ 0 & -1 & 2+h^2q_3 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 2+h^2q_{n-1} \end{pmatrix}}_{A \in \mathbb{R}^{(n-1)\times(n-1)}} \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ ... \\ y_{n-1} \end{pmatrix}}_{y} = \underbrace{\begin{pmatrix} h^2g_1 + \alpha \\ h^2g_2 \\ h^2g_3 \\ ... \\ h^2g_{n-1} + \beta \end{pmatrix}}_{b}$$

Fragen:

1. Ist das GS eindeutig lösbar?
2. Effiziente Lösungsverfahren?
3. Konvergenz der num. Approximation gegen exakte Lösungen für $h \to 0$?

zu 1):
Eigenschaften von A:

- symmetrisch, tridiagonal
- positiv definit, da:
  $A = D + T;\ T = tridiag(-1, 2, -1);\ D = diag(h^2 q_i,\ i = 1, ..., n-1)$
  $\Rightarrow \forall x \in \mathbb{R}^{n-1}$:
  $(Ax, x) = (Tx, x) + h^2 \sum_{t=1}^{n-1} \underbrace{q_i}_{\geq 0} \geq (Tx, x)$
  Z.z.: T positiv definit, d.h., $(Tx, x) > 0 \forall x \neq 0$.
  Matrix T besitzt folgende Eigenwerte und Eigenvektoren:
  Eigenwerte von T: $\lambda_k = 4 \sin^2 \frac{\pi k}{2n} > 0,\ k = 1, ..., n-1$
  Eigenvektoren: $(v_k)_j = \sin \frac{\pi k j}{n}$

- Man kann zeigen, daß für $2 \leq j \leq n-2$ gilt:

$$\underbrace{-\sin\frac{\pi k(j-1)}{n}}_{(v_k)_{j-1}} + \underbrace{2\sin\frac{\pi kj}{n}}_{=(v_k)_j} - \sin\frac{\pi k(j+1)}{n} \overset{Additionstheoreme}{=} \underbrace{4\sin^2\frac{\pi k}{2n}}_{=\lambda_k} \cdot \sin\frac{\pi kj}{n},$$

  was zur eindeutige Lösbarkeit des Diskretisierungssystems ($Ay = b$) führt.

zu 2): A ist (strikt) diagonaldominant (da $q_j \leq 0$) $\Rightarrow O(n)$-Eliminationsalgorithmus (z.B. Thomas-Algorithmus, Gauß-Algorithmus für tridiagonale Matrizen) ist stets durchfürbar und stabil.

6 Anhang

6.1 Landau-Symbole

(größtenteils basiert an der Vorlesungen von Prof. Weikert)

- Suchen kompakte Notation zur Aufwandsabschätzung von Algorithmen.
- Algorithmus soll von Problemgröße n abhängen (z.B.: n = Zahl der zu sortierenden Objekte in einer Liste)
- Laufzeit ist Abbildung $f : \mathbb{N} \to \mathbb{R}^+$ mit $f(n) = a_n$, d.h. eine Folge.
- Meist ist (a_n) monoton wachsend und nicht beschränkt, d.h. $lim_{n\to\infty} a_n = 1$.
- Interessant ist, wie schnell a_n gegen ∞ strebt. Hierzu vergleicht man (a_n) mit Prototypen von anderen Folgen (b_n).

Definition 55 *(Groß-O,Klein-o)*
Ein Folge $A = (a_n)$ ist "Groß-O" von $B = (b_n)$, wenn der Quotient $\left(\frac{a_n}{b_n}\right)$ beschränkt ist. Die Folge A ist "Klein-o" von B, wenn $\left(\frac{a_n}{b_n}\right)$ eine Nullfolge ist. Man schreibt $A = O(B)$ bzw. $A = o(B)$. O und o nennt man auch Landau-Symbole.

Beispiel 56 *1.* $2n^2 + 3n + 4 = O(n^2)$, *denn*

$\frac{2n^2+3n+4}{n^2} = \frac{n^2(2+\frac{3}{n}+\frac{4}{n^2})}{n^2} = 2 + \frac{3}{n} + \frac{4}{n^2}$

Dies ist eine konvergente Folge und damit beschränkt.

2. $2n^2 + 3n + 4 = o(n^3)$, *denn*

$\frac{2n^2+3n+4}{n^3} = \frac{\frac{2}{n}+\frac{3}{n^2}+\frac{4}{n^2}}{1} \to 0$ *für* $n \to \infty$.

Bemerkung 57 ***Mehrdeutigkeit***
Die Angabe $A = O(B)$ ist nicht eindeutig. Für $A = (n + 2)$ sind z.B.:

$$A = O(n^2)$$

$$A = O(n)$$

$$A = O(\frac{n}{2})$$

wahre Aussagen. Es gilt:

- *Konstante Faktoren ändern nicht die Ordnung.*

$$O(B) = O(c \cdot B) \quad \forall c \in \mathbb{R}.$$

- *Terme niedrieger Ordnung sind unwichtig:*
  z.B. haben $(n^2 + 3n + 4)$ *und* $(n^2 + 17n)$ *die selbe Ordnung* $O(n^2)$.

Satz 58 *(Rechenregeln für Landau-Symbole)*
Für reelle Zahlenfolgen A, B *und* C *gelten folgende Regeln:*

1. $A = O(A)$
2. $\forall c \in \mathbb{R}$ *gilt:*
   $c \cdot O(A) = O(A)$
   $c \cdot o(A) = o(A)$
3. $O(A) + O(A) = O(A)$
   $o(A) + o(A) = o(A)$
4. $O(O(A)) = O(A)$
   $o(o(A)) = o(A)$
5. $O(A) \cdot O(B) = O(A \cdot B)$
   $o(A) \cdot o(B) = o(A \cdot B)$
6. $A \cdot O(B) = O(A \cdot B)$
   $A \cdot o(B) = o(A \cdot B)$
7. *Transitivität:*
   $A = O(B), \quad B = O(C) \Rightarrow A = O(C)$
   $A = o(B), \quad B = o(C) \Rightarrow A = o(C)$

Bemerkung 59 ***Vergleichbarkeit von Folgen***
Nicht alle Folgen sind vergleichbar!
Z.B. $A = (a_n)$ *und* $B = (b_n)$ *mit*

$$a_n = \begin{cases} n & (n \text{ gerade}) \\ n^2 & (n \text{ ungerade}) \end{cases}$$

$$b_n = \begin{cases} n^2 & (n \text{ gerade}) \\ n & (n \text{ ungerade}) \end{cases}$$

Für gerades n *sieht man, daß* $A \neq O(B)$.
Für ungerades n *sieht man , daß* $B \neq O(A)$.

Definition 60 *(*$O(A) = O(B),\ O(A) < O(B)$*)*
Man definiert folgende Relationen:

$$O(A) = O(B) \quad :\Leftrightarrow \quad A = O(B) \text{ und } B = O(A)$$
$$O(A) < O(B) \quad :\Leftrightarrow \quad A = O(B) \text{ und } B \neq O(A)$$

Bemerkung 61 ***Häufig verwendete Prototypen von Vergleichsfunktionen***

O	*Laufzeitverhalten*
$O(1)$	*konstant*
$O(\log_a n),\ a > 1$	*logarithmisch*
$O(n)$	*linear*
$O(n \log_a n),\ a > 1$	$n \log n$
$O(n^2)$	*quadratisch*
$O(n^3)$	*kubisch*
$O(n^k)$	*polynomial*
$O(a^n),\ a > 1$	*exponentiell*

Häufig ist $a = 2$. Die Wahl von a ist jedoch unwichtig, da sich Logarithmen zu unterschiedlichen Basen nur durch einen konstanten Faktor unterscheiden:

$$\log_a n = \frac{\log_c n}{\log_c a}.$$

Bemerkung 62 ***Vergleich von Ordnungen***

Es gilt:

$$\begin{array}{ccccccccccccccc} O(1) & < & O(\log n) & < & O(n) & < & O(n \log n) & < & O(n^2) & < & O(n^3) & < & \\ & < & \ldots & < & O(n^k) & < & \ldots & < & O(2^n) & < & O(3^n) & < & \ldots \end{array}$$

Z.B.:

n	$\log_2 n$	$n \log_2 n$	n^2	n^3	2^n
10	3.32	33.32	100	1000	1024
100	6.64	664.4	10000	10^6	$1.27 \cdot 10^{30}$
1000	9.97	9966	10^6	10^9	$\approx 10^{301}$
10000	13.29	132877	10^8	10^{12}	$\approx 10^{3010}$

Für große n sollte man daher nicht auf Fortschritte im Rechnerbau hoffen, sondern nach Algorithmen suchen, die eine bessere (also: kleinere) Ordnung besitzen.

6.2 Eigenwerte und Eigenvektoren

Definition 63 (Eigenvektor, Eigenwert) *Sei $f : V \to V$ eine lineare Abbildung von K-Vektorräumen. Ein Vektor $x \neq 0$ aus V heisst Eigenvektor von f, wenn es $\lambda \in K$ gibt mit $f(v) = \lambda v$. Die Zahl $\lambda \in K$ heisst dann der zugehörige Eigenwert von f.*

Beispiel 64 *Sei* $A := \begin{pmatrix} 2 & -1 & 0 \\ 2 & 2 & 2 \\ 2 & 3 & 1 \end{pmatrix}$ *und* $v := \begin{pmatrix} 1 \\ 3 \\ -\frac{11}{2} \end{pmatrix}$ *Dann rechnet man leicht nach:*

$$Av = \begin{pmatrix} -1 \\ -3 \\ -\frac{11}{2} \end{pmatrix}$$

Also ist v eine Eigenvektor zum Eigenwert $\lambda = -1$.

Definition 65 *Sei $f : V \to V$ eine lineare Abbildung von K-Vektorräumen. Dann heisst der Unterraum*

$$E_\lambda := \{v \in V | fv = \lambda v\}$$

Der Eigenraum von f zum Eigenwert λ. Er enthält genau die Eigenvektoren zum Eigenwert λ und den Nullvektor.

Bemerkung 66 *Einen Spezialfall von Eigenräumen kennen wir schon länger:*

$$E_0 = \ker f$$

Definition 67 (charakteristisches Polynom) *Sei $A \in M(n \times n, K)$ Dann ist das charakteristische Polynom gegeben als*

$$\chi(A) = \det(XE_n - A)$$

Satz 68 *λ Nullstelle von $\chi(A)$ gdw. λ Eigenwert von A.*

Beweis: „⇒ “ Sei $\det(\lambda E_n - A) = 0$. Dann ist $\lambda E_n - A$ singulär, und hat nichttrivialen Kern. Sei v im Kern, dann gilt:

$$(\lambda E_n - A)v = 0 \Rightarrow \lambda v = Av$$

„⇐ “Alle Implikationen sind umkehrbar.

Aufgaben 69 *Bestimme die Eigenwerte der folgenden Matrizen:*

a) $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 4 & 3 \\ 1 & -4 & -5 \end{pmatrix}$ *(Lösung:* $0, 2\sqrt{3}, -2\sqrt{3}$*)*

b) $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ *(Lösung:* $2, i, -i$*)*

Definition 70 *Eine quadratische Matrix A heisst diagonalisierbar, wenn eine invertierbare Matrix P existiert, so dass $P^{-1}AP$ Diagonalgestalt hat.*

Satz 71 *$A \in K^{n\times n}$ ist diagonalisierbar gdw. es ex. n lin. unabh. Eigenvektoren von A. In diesem Fall sind die Spalten von P Eigenvektoren von A.*

Aufgaben 72 *Man untersuche jeweils, ob die Matrix A über K diagonalisierbar ist. Wenn ja, so gebe man eine invertierbare Matrix P und eine Diagonalmatrix D an mit $D = P^{-1}AP$.*

1. $K = \mathbb{R}$, $A = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$
2. $K = \mathbb{C}$, $A = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}$

Lösung

(a)

$$\chi_A(x) = (2-x)^2 + 1 = x^2 - 4x + 5$$

Das charakteristische Polynom hat keine reellen Nullstellen, zerfällt also nicht und ist somit nicht diagonalisierbar

(b)

$$\begin{aligned} \chi_A(x) &= (5-x)(4-x)(-4-x) - 36 - 36 + 18(4-x) + 12(5-x) - 6(-4-x) \\ &= -x^3 + 5x^2 - 8x + 4 \\ &= (1-x)(2-x)^2 \end{aligned}$$

A hat somit 2 und als 1 als Eigenwerte.Man berechnet einfach die Eigenräume:
Die Eigenvektoren zum Eigenwert 2 ergeben sich durch Lösung von

$$(A - 2I)X = 0$$

Man erhält $(2,1,0), (2,0,1)$ und zum Eigenwert 1 den Eigenvektor $(3,-1,3)$ Somit ist

$$P = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 0 & -1 \\ 0 & 1 & 3 \end{pmatrix}$$

Satz 73 *Sei A eine Matrix in Dreiecksform. Dann sind die Diagonaleinträge genau die Eigenwerte der Matrix.*

Aufgaben 74 *1. Sei $A \in \mathbb{C}^{n*n}$. Zeige: Für eine beliebeige $n*n$ Matrix B und jeden Eigenwert λ von A gilt die Ungleichung*

$$1 \leq ||(\lambda I - B)^{-1}(A - B)|| \leq ||(\lambda I - B)^{-1}||\ ||A - B||$$

sofern λ nicht auch Eigenwert von B ist. Hierbei bezeichnet I die $n \times n$-Einheitsmatrix und $||\ ||$ eine beliebige zugeordnete (also insbesondere auch submultiplikative) Matrixnorm.

*2. Beweise **Satz von Gerschgorin:***

Sei $A = (a_{ij})_{i,j=1}^n$ eine $n \times n$ Matrix. Dann liegen alle Eigenwerte der Matrix in der Menge

$$M := \bigcup_{i=1}^{n} K_i, \quad K_i := \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{j=1, j\neq i}^{n} |a_{ij}| \right\} (i = 1, ..., n).$$

Lösung

1. Sei x EV zu EW λ von A, λ kein EW von B. Dann ist $\lambda I - B$ regulär.

$$(A - B)x = (\lambda I - B)x$$
$$\Rightarrow (\lambda I - B)^{-1}(A - B)x = x$$

   $||\ ||$ Vektornorm $\Rightarrow$

$$||(\lambda I - B)^{-1}(A - B)x|| = ||x||$$

   Laut Aufgabenstellung ist die Norm submultiplikativ, d.h.

$$||A \cdot B|| \leq ||A|| \cdot ||B||$$

   (Bemerkung: Das gilt ohnehin für jede von einer Vektornorm induzierten Norm)
   Damit gilt:

   $\Rightarrow ||(\lambda I - B)^{-1}||\ ||A - B||\ ||x|| \geq ||(\lambda I - B)^{-1}(A - B)||\ ||x|| \geq ||x||$

   Dinvision durch $||x||$ liefert die Behauptung.

2. Ist ein a_{ii} EW von $A \Rightarrow$ Behauptung.

   Sonst:

   verwende 1. mit $B = diag(a_{11}, ..., a_{nn})$ und der $||\ ||_\infty$-Norm

$$\max_{1 \leq i \leq n} |\lambda - a_{ii}|^{-1} \sum_{j \neq i} |a_{ij}| \geq 1$$

   d.h. exisitiert i_0 mit $|\lambda - a_{i_0 i_0}| \leq \sum_{j \neq i_0} |a_{i_0 j}|$

   Dann liegt λ auf jeden Fall in M.

   λ beliebig $\Rightarrow$ Behauptung.

Bemerkung 75 *Es kann Fälle geben, in denen eine (oder sogar mehere-re) der Mengen K_i keinene Eigenwert erhält, z.B.:*
$A = \begin{pmatrix} 0 & -1 \\ 4 & -4 \end{pmatrix}$ *hat doppelten EW* -2.
$K_1 = \{|z| < 1\}$ *enthält EW nicht.*

6.3 Orthogonale Matrizen

Definition 76 *Eine Matrix $A \in K^{n\times n}$ heisst orthogonal, wenn ihre Spalten eine Orthonormalbasis des K^n bilden.*
Die Menge der orthogonalen $n \times x$-Matrizen heisst $O(n)$.

Satz 77 *Sei $A \in O(n)$ Dann gilt:*

1. $$A^{-1} = A^T$$
2. $$\langle Au, Av\rangle = \langle u, v\rangle \forall u, v \in \mathbb{R}^n$$
   (A ist Isometrie)
3. $$\| Au \| = \| u \| \ \forall u \in \mathbb{R}^n$$
4. *Orthogonale Matrizen haben immer Determinante +1 oder -1*
5. *Nur ± 1 kommen als Eigenwerte von A in Frage.*
6. *O(n) ist eine Gruppe bzgl. Matizenmultiplikation. Die orthogonalen Matrizen mit Determinante +1 bilden eine Untergruppe von $O(n)$, die spezielle orthogonale Gruppe:*
   $$SO(n) := \{A \in O(n); \det(A) = +1\}$$

Bemerkung 78 *Vorsicht! Nicht jede Matrix mit Determinante ± 1 ist auch orthogonal:*
$$\begin{pmatrix} 3 & 3 \\ 2 & 4 \end{pmatrix}$$

Definition 79 *Eine lineare Abbildung $F : \mathbb{R}^n \to \mathbb{R}^n$ heisst orthogonal, wenn ihre Matrix bzgl. der Standardbasis orthogonal ist (äquivalent: Wenn ihre Matrix bzgl einer ON-Basis von $\mathbb{R}^n$ orthogonal ist).*

Beispiel 80 *Die Elemente in O(2) haben die Form*
$$\begin{pmatrix} \cos\alpha & -sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \text{ oder } \begin{pmatrix} \cos\alpha & sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$
mit $\alpha \in [0, 2\pi[$

Aufgaben 81 *Zeige: $F : \mathbb{R}^3 \to \mathbb{R}^3$ speziell orthogonal. $\Rightarrow$ F besitzt einen Eigenvektor zum Eigenwert* 1.

Lösung: Sei F orthogonal, A seine Matrix bzgl. der Standardbasis. Dann ist das charakteristische Polynom von A ein Polynom dritten Grades, hat also mind. eine reelle Nullstelle.
Es gibt also einen Eigenwert $\lambda_1 = \pm 1$. Ist $\lambda_1 = 1$, so sind wir fertig. Ansonsten könnnen wir eine ON-Basis $\{w_1, w_2, w_3\}$ wählen, bzgl der F die Darstellung

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}$$

hat. Sei

$$A' := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Dann gilt $A' \in O(2)$ und $\det A' = -1$, also ist

$$A' = \begin{pmatrix} \cos\alpha & sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

Das charakteristische Polynom ist $t^2 - 1$, also exisitiert ein Eigenvektor zum Eigenwert 1.

Geometrisch beschreiben die speziell orthogonal Abbildungen im $\mathbb{R}^3$ Drehungen im Raum. Wir können also folgern:

Korollar 82 (Satz vom Fussball) *Auf der Oberfläche eines Fussballs gibt es einen Punkt, der sich beim Anstoss zur ersten und zweiten Halbzeit an der gleichen Stelle befindet.*

Beweis: Da der Ball an der gleichen Stelle liegt (auf dem Anstosspunkt), wird seine „Positionsänderung“ durch eine Drehung im $\mathbb{R}^3$ beschrieben, wobei der Mittelpunkt des Balles der Nullpunkt ist. Nach vorherigem Satz gibt es einen Eigenvektor v zum Eigenwert 1, dieser entspricht genau einem Punkt auf der Oberfläche des Balls. Dieser wird durch die Drehung unverändert gelassen (das bedeutet Eigenwert 1).

6.4 Symmetrische Matrizen

Satz 83 (Eigenschaften symmetrischer Matrizen) *Für eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ gilt*

1. *A hat nur reelle Eigenwerte.*
2. *Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.*

6.5 Numerische Differentiation (Finite Differenzen)

(größtenteils basiert an der Vorlesungen von Prof. Weikert)

In vielen Fällen muss man Ableitungen nur diskret vorliegenden Funktionen (d.h. nur an einzelnen Punkten, 'Gitterfunktionen', s. z.B. Kapitel 5.3) approximieren.

In der Praxis treten solche Tabellenfunktionen, z.B., bei Geschwindigkeitsmessung durch die Polizei oder bei Temperaturmessungen auf.

Man wendet den Satz von Taylor an, um geeignete numerische Approximationsformeln zu bestimmen.

Problembeschreibung:

Von einer kontinuierlichen Funktion $f(x)$ sind nur die Funktionswerte ('Messwerte', 'Stützwerte') auf einem äquidistanten Gitter $G := \{x_i | x_i = ih, i \in \mathbb{Z}\}$ ('Stützstellen') bekannt:

$$f_i = f(x_i),\ i \in \mathbb{Z}.$$

h heißt die Gitterweite von G.

Gesucht wird Approximation einer Ableitung $f^{(k)}$ an der Stelle ξ, die nicht notwendigerweise an dem Gitter liegt, so daß nur die bekannte Größen f_i verwendet werden: Finite-Differenzen-Approximation.

Beispiel 84 *Approximieren wir f'' im Punkt x_i mit Hilfe von f_{i-1}, f_i, f_{i+1}. Es werden also Zahlen α, β, γ gesucht, so daß*

$$f_i'' \approx \alpha f_{i-1} + \beta f_i + \gamma f_{i+1}$$

gilt, und Approximationsfehler dabei höchstens $O(h)$ ist.

Lösung:

Sei f hinreichend oft differenzierbar.

Laut Aufgabestellung gilt:

$$f_i'' = \alpha f_{i-1} + \beta f_i + \gamma f_{i+1} + O(h).$$

Mit der Taylorentwicklung von f um x_i gilt:

$$\begin{aligned} \alpha f_{i-1} + \beta f_i + \gamma f_{i+1} &= \\ &= \alpha(f_i - hf_i' + \tfrac{1}{2}h^2 f_i'' + O(h^3)) + \\ &\quad +\beta f_i + \\ &\quad +\gamma(f_i + hf_i' + \tfrac{1}{2}h^2 f_i'' + O(h^3)) = \\ &= f_i(\alpha + \beta + \gamma) + f_i'(-h\alpha + h\gamma) + f_i''\tfrac{h^2}{2}(\alpha + \gamma) + O(h^3) \end{aligned}$$

Da dies mit $f_i'' = 0 \cdot f_i + 0 \cdot f_i' + 1 \cdot f_i''$ übereinstimmen soll, liefert ein Koeffizientenvergleich das lineare GS:

$$\begin{array}{llcll} (1) & \alpha + \beta + \gamma & = & 0 & \\ (2) & -h\alpha + h\gamma & = & 0 & \Rightarrow \quad \alpha = \gamma \\ (3) & \frac{h^2}{2}\alpha + \frac{h^2}{2}\gamma & = & 1 & \Rightarrow \quad h^2\alpha = 1 \end{array}$$

Aus (2) folgt $\alpha = \gamma$.
Zusammen mit (3): $h^2\alpha = 1 \Rightarrow \alpha = \gamma = 1/h^2$.
Aus (1) folgt damit: $\beta = 1 - \alpha - \gamma = -2/h^2$.
Also gilt:

$$f_i'' \approx \frac{f_{i-1} - 2f_i + f_{i+1}}{h^2}.$$

Wie genau ist diese Approximation?

Mit der ausführlichen Taylorentwicklung

$$f_{u\pm i} = f_i \pm h f_i' + \frac{1}{2}h^2 f_i'' \pm \frac{1}{6}h^3 f_i''' + \frac{1}{24}h^4 f_i^{(4)} \pm \frac{1}{120}h^5 f_i^{(5)} + \frac{1}{720}h^6 f_i^{(6)} + O(h^7)$$

folgt durch Einsetzen:

$$\frac{f_{i-1} - 2f_i + f_{i+1}}{h^2} = f_i'' + \underbrace{\frac{1}{12}h^2 f_i^{(4)}}_{führender\ Fehlerterm} + O(h^4)$$

Der führender Fehlerterm hat also die Ordnung $O(h^2)$. *Man sagt, die Approximation*

$$f_i'' \approx \frac{f_{i-1} - 2f_i + f_{i+1}}{h^2}$$

hat die Konsistenszordnung $O(h^2)$ *(oder einfach Konsistenzordnung 2):*

$$f_i'' = \frac{f_{i-1} - 2f_i + f_{i+1}}{h^2} + O(h^2).$$

Anmerkungen:

a) Höhere Konsistenzordnung bedeutet bessere Approsimation für $h \to 0$. Die Minimalanforderung lautet 'mindestens lineare Konsistenzordnung $O(h)$' - dann heißt die Approximation konsistent.

b) Die Konsistenzordnung hängt vom Entwicklungspunkt ab, z.B.:

- Approximiert man den Wert von f_i' mit f_i, f_{i+1}, so folgt mit der TAylorentwicklulng um x_i

  $$f_{i+1} = f_i + h f_i' + 1/2 h^2 f_i'' + O(h^3)$$

  die Approximation der Ordnung $O(h)$:

  $$f_i' = \frac{f_{i+1} - f_i}{h} + 1/2 h f_i'' + O(h^2).$$

  Entwickelt man jedoch um $x_{i+1/2}$:

  $$f_{i+1} = f_{i+1/2} + \frac{h}{2}f_{i+1/2}' + \frac{1}{2}\frac{h^2}{4}f_{i+1/2}'' + \frac{1}{6}\frac{h^3}{8}f_{i+1/2}''' + O(h^4)$$

$$f_i = f_{i+1/2} - \frac{h}{2} f'_{i+1/2} + \frac{1}{2}\frac{h^2}{4} f''_{i+1/2} - \frac{1}{6}\frac{h^3}{8} f'''_{i+1/2} + O(h^4),$$

ergibt sich eine $O(h^2)$-Approximation:

$$f_{i+1/2} = \frac{f_{i+1} - f_i}{h} + \frac{1}{24} h^2 f''_{i+1/2} + O(h^4).$$

c) Durch hinzunahme weiterer Gitterpunkte gelingt es oft, die Konsistenzordnung zu verbessern, z.B.:

$$f'_i = \frac{f_{i+1} - f_{i-1}}{2h} + O(h^2)$$

$$f'_i = \frac{-f_{i+2} + 8f_{i+1} - 8f_{i-1} + f_{i-2}}{12h} + O(h^4).$$

d) Wichtige Approximationen:

Erste Ableitung:

$$\begin{array}{ll} f'_i = \frac{f_{i+1}-f_i}{h} + O(h) & Vorwärtsdifferenz \\ f'_i = \frac{f_i-f_{i-1}}{h} + O(h) & Rückwärtsdifferenz \\ f'_i = \frac{f_{i+1}-f_{i-1}}{h} + O(h^2) & zentrale\ Differenz \end{array} \tag{7}$$

Zweite Ableitung:

$$f''_i = \frac{f_{i+1} - 2f_i + f_{i-1}}{h^2} + O(h^2).$$

e) Will man partielle Ableitungen einer Funktion mehrerer Variablen approximieren, benotigt man oft eine Taylorentwicklung in mehreren Variablen. Z.B.:

$$\left.\frac{\partial^2 f}{\partial x \partial y}\right|_{i,j} \approx \frac{1}{4hk}[f_{i+1,j+1} + f_{i-1,j-1} - f_{i+1,j-1} - f_{i-1,j+1}]$$

mit $f_{i,j} = f(x_i, y_j)$ und $x_i = ih,\ y_j = jk$, wobei $h,\ k$ Gitterweiten entlang $x-$ und $y-$Achsen sind.

Manchmal hat man aber das Glück, daß eindimensionale Taylorentwicklungen längs der Koordinatenachsen ausreichen, z.B.:

$$\Delta f|_{i,j} = \frac{\partial^2 f}{\partial x^2} + \left.\frac{\partial^2 f}{\partial y^2}\right|_{i,j} = \frac{f_{i+1,j} - 2f_{i,j} + f_{i-1,j}}{h^2} + \frac{f_{i,j+1} - 2f_{i,j} + f_{i,j-1}}{h^2} + O(h^2 + k^2).$$

Anwendung auf die Diskretisierung von Differentialgleichungen:

Ersetzt man die Ableitungen durch Differenzenapproxiamtionen, entstehen so genannte Finite-Differenzen-Verfahren.

Beispiel 85 *Die eindimensionale Diffusionsgleichung*

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

soll mit einema einfachen finiteDifferenzen-Verfahren numerisch approximiert werden. Bezeichnet u_i^k eine Approxiamtion an $u(x,t)$ im Punkt $(ih, k\tau)$, kann man beispielsweise

$$\left.\frac{\partial u}{\partial t}\right|_i^k = \frac{u_i^{k+1} - u_i^k}{\tau} + O(\tau)$$

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_i^k = \frac{u_{i+1}^k - 2u_i^k + u_{i-1}^k}{h^2} + O(h^2)$$

verwenden. damit lautet die Approximation

$$\frac{u_i^{k+1} - u_i^k}{\tau} = \frac{u_{i+1}^k - 2u_i^k + u_{i-1}^k}{h^2}.$$

Kennt man die Funktionswerte zur Zeitschicht $k = 0$, kann man damit die unbekannten Werte der Schichten $k = 1, 2, ...$ iterativ berechnen:
Für $k = 0, 1, ...$

$$\begin{aligned} u_i^{k+1} &= u_i^k + \tfrac{\tau}{h^2}(u_{i-1} - 2u_i^2 + u_{i+1}) \\ &= \tfrac{\tau}{h^2} u_{i+1}^k + \left(1 - 2\tfrac{\tau}{h^2}\right) u_i^k + \tfrac{\tau}{h^2} u_{i-1}^k. \end{aligned}$$

u_i^{k+1} entsteht somit durch gewichtete Mittelung aus u_{i-1}^k, u_i^k und u_{i+1}^k.