Technische Universität München
Institut für Informatik
Univ.-Prof. Dr. Hans-Joachim Bungartz
Michael Obersteiner
Philipp Samfass

WiSe 2017 / 2018

Numerisches Programmieren, Übungen

Musterlösung 10. Übungsblatt: Gewöhnliche Differentialgleichungen (ODE) II

1) Begriffe

Was versteht man unter den folgenden Begriffen:

a) lokaler Diskretisierungsfehler,

b) globaler Diskretisierungsfehler,

c) Konvergenz,

d) Konsistenz,

e) Stabilität,

f) Steifheit?

Machen Sie sich insbesondere den Unterschied zwischen lokalem und globalem Diskretisierungsfehler klar. Nutzen Sie hierfür eine Skizze!

Lösung:

a) **lokaler Diskretisierungsfehler:**

Größter Fehler, der durch die Verwendung des Differenzenquotentienten anstatt der Ableitung entsteht.
Lokaler Diskretisierungsfehler der Eulermethode:

$$l(\delta t) \ := \ \max_{a \leq t \leq b-\delta t} \left\{ \left| \frac{y(t+\delta t) - y(t)}{\delta t} - f(t, y(t)) \right| \right\}$$

Oft wird anstatt dieser eher unhandlichen Definition des lokalen Diskretisierungsfehlers (der Eulermethode) eine andere (nicht äquivalente) Darstellung verwendet, die in der Praxis leichter handhabbar ist. Sie verwendet als lokalen Diskretisierungsfehler die Differenz zwischen analytischer und numerischer Lösung am Ende eines einzelnen Zeitschritts:

$$|y_{k+1} - y(t_{k+1})|,$$

wobei $y_k = y(t_k)$ gilt.

b) **globaler Diskretisierungsfehler:**

Größter Fehler zwischen den numerischen Approximationen y_k und den korrespondierenden analystischen Werten $y(t_k)$ zu konkreten Zeitpunkten t_k:

$$e(\delta t) \ := \ \max_{k=0,\ldots,N} \{|y_k - y(t_k)|\}$$

Eine in der Praxis eher gebräuchliche (aber nicht äquivalente) Form des globalen Diskretisierungsfehlers betrachtet lediglich die Differenz zwischen numerischer und analytischer Lösung der Differentialgleichung am Ende des betrachteten Zeitintervalls (nach n Zeitschritten zum Endzeitpunkt $t = b$):

$$|y_n - y(t_n = b)|$$

wobei $y_0 = y(t_0)$ gilt.

c) **Konvergenz:**

Konvergiert die durch ein Lösungsverfahren erzeugte Folge mit immer kleiner werdender Schrittweite ($h \to 0$) gegen die exakte Lösung, so heißt dieses Lösungsverfahren *konvergent*.
Es gilt also:

$$\lim_{h \to 0, n \to \infty} y_k = y(t_k)$$

Der globale Fehler konvergiert folglich gegen 0.

d) **Konsistenz:**

Der Begriff der *Konsistenz* ist eng mit dem lokalen Diskretisierungsfehler verbunden. Falls für $\delta t \to 0$ der lokale Fehler $l(\delta t)$ gegen 0 konvergiert (also $l(\delta t) \to 0$), so heißt ein Verfahren *konsistent*.

e) **Stabilität:**

- *Stabilität* ist die Eigenschaft eines Verfahrens/Algorithmus.
- Ein Verfahren ist stabil, wenn es gegenüber kleinen Störungen unempfindlich ist.
- Summieren sich kleine lokale Fehler nur zu kleinen globalen Fehlern auf, so ist ein Verfahren stabil.

f) **Steifheit:**

Die Begriffe Konsistenz und Konvergenz sind eher analytischer Natur: Für ausreichend/beliebig kleine δt soll etwas passieren. Es gibt jedoch Differentialgleichungen, die konsistent bzw. konvergent sind, jedoch nur für sehr kleine δt. Ein zu kleines δt kann die Diskretisierung jedoch unpraktikabel machen.
Differentialgleichungen mit dieser Eigenschaft bezeichnet man als *steif*.
Somit ist Steifheit eine Problemeigenschaft, die bei der numerischen Lösung der Differentialgleichung mit expliziten Einschrittverfahren die Verwendung einer sehr kleinen Schrittweite δt erzwingt, die für eine vorgegebene Genauigkeit in der Approximation eigentlich nicht nötig wäre.

Für die Beziehung zwischen Stabilität, Konsitenz und Konvergenz gilt bei den expliziten Verfahren:

- ESV: Konsistenz $\Rightarrow$ Konvergenz
- MSV: Konsistenz + Stabilität $\Leftrightarrow$ Konvergenz

2) Quadratur und AWP-Lösung

In dieser Aufgabe wollen wir uns noch einmal die Analogie von Quadratur und numerischer Lösung von Anfangswertproblemen (AWP) verdeutlichen. Wie in der Vorlesung definieren wir ein AWP durch eine gewöhnliche Differentialgleichung (1) und einen Anfangswert (2)

$$\dot{y}(t) = f(t, y(t)) \tag{1}$$
$$y(a) = y_0. \tag{2}$$

Ziel ist die Approximation y_i der zugehörigen Lösungsfunktion $y(t_i)$ zu bestimmten Zeitpunkten t_i. Wenn wir **Ein**schrittverfahren verwenden, interessiert uns immer ein Teilintervall $[t_k; t_{k+1}]$, um aus **einem** alten Wert y_k den neuen y_{k+1} zu berechnen. Dabei können wir den Hauptsatz der Differential- und Integralrechnung benutzen, der uns folgenden Zusammenhang liefert:

$$y(t_{k+1}) - y(t_k) = \int_{t_k}^{t_{k+1}} \dot{y}(t)\, dt \overset{(1)}{=} \int_{t_k}^{t_{k+1}} f(t, y(t))\, dt\ . \tag{3}$$

Um nun approximierte Werte y_k und y_{k+1} auf der linken Seite von (3) zu erhalten, integriert man die rechte Seite numerisch.

a) Benutzen Sie die im Folgenden definierte "dumme Rechtecksregel", um aus (3) das explizite Euler-Verfahren herzuleiten!
Die "dumme Rechtecksregel" arbeitet wie die normale Rechtecksregel, nur wird die Funktion am linken Rand des Intervalls ausgewertet und nicht in der Mitte:

$$\int_a^b f(t)\, dt \approx (b-a) \cdot f(a)\ .$$

b) Benutzen Sie die Trapezregel sowie einen zusätzlichen Approximationsschritt, um aus (3) das Verfahren von Heun herzuleiten!

Lösung:

a) Die "dumme Rechtecksregel"

$$\int_a^b f(t)\, dt \approx (b-a) \cdot f(a)\ .$$

liefert direkt

$$\begin{aligned} y_{k+1} - y_k &= \underbrace{(t_{k+1} - t_k)}_{=\delta t} \cdot f(t_k, y_k) \\ \Rightarrow \quad y_{k+1} &= y_k + \delta t \cdot f(t_k, y_k)\ . \end{aligned}$$

b) Die Trapezregel lautet

$$\int_a^b f(t)\, dt \approx (b-a) \cdot \frac{1}{2}\left[f(a) + f(b)\right]\ .$$

Damit ergibt sich zunächst

$$\begin{aligned} y_{k+1} - y_k &= \underbrace{(t_{k+1} - t_k)}_{=\delta t} \cdot \frac{1}{2} \left[f(t_k, y_k) + f(t_{k+1}, y_{k+1}) \right] \\ \Rightarrow \quad y_{k+1} &= y_k + \delta t \cdot \frac{1}{2} \left[f(t_k, y_k) + f(t_{k+1}, y_{k+1}) \right] \; . \end{aligned} \tag{4}$$

Allerdings taucht in Glg. (4) nun auch auf der rechten Seite das noch unbekannte y_{k+1} auf. Um die Lösung einer (i.A.) nichtlinearen Gleichung zu vermeiden, approximiert man diesen Wert mit Hilfe eines einfach zu berechnenden anderen Verfahrens, wie z. B. des expliziten Euler-Verfahrens. So erhalten wir:

$$y_{k+1} = y_k + \delta t \cdot \frac{1}{2} \left[f(t_k, y_k) + f(t_{k+1}, \underbrace{y_k + \delta t \cdot f(t_k, y_k)}_{\approx y_{k+1}}) \right] \; . \tag{5}$$

3) Euler-Verfahren und Zinsberechnung

Die Verzinsung eines Guthabens, das am Anfang den Wert y_0 habe und pro Jahr $p\%$ Zinsen erhält, kann man als Anwendung des expliziten Euler-Verfahrens auf ein Anfangswertproblem $\dot{y} = f(y)$ interpretieren.

a) Wiederholung: Wie sieht die Berechnungsvorschrift des expliziten Euler-Verfahrens mit Schrittweite δt für das AWP $\dot{y} = f(y)$ aus?

b) Die Diga-Bank schüttet Zinsen immer zum Jahresende aus. Geben Sie die rechte Seite $f(y)$ an, so dass das Eulerverfahren mit δt=1 (Jahr) gerade diese Verzinsung beschreibt.

c) Berechnen Sie die analytische Lösung des AWPs!

d) Nun macht die Spaßkasse das Angebot, Zinsausschüttungen vierteljährlich statt nur am Ende des Jahres durchzuführen. Geben Sie die Euler-Formel für diese neue Schrittweite an und berechnen Sie daraus eine explizite Vorschrift, um direkt den Wert des Gesamtguthabens nach einem vollen Jahr (y_{end}) zu ermitteln.

e) Vergleichen Sie für den konkreten Fall eines Startguthabens von y_0 =100000 Euro und einer Verzinsung von $p = 2$ Prozent p.a. die verschiedenen Guthaben y_{end} nach einem Jahr bei der Diga-Bank, der Spaßkasse und einer virtuellen Bank mit Verzinsung nach der analytischen Lösung des AWPs! Was stellen Sie fest?

Lösung:

a) Explizites Euler-Verfahren für das angegebene AWP:

$$y_{k+1} = y_k + \delta t \cdot f(y_k) \; .$$

b) Ausgehend vom Startkapital y_0 soll mit $\delta t = 1$ das Startkapital plus Zinsen bei y_1 herauskommen. Die Zinsen für den Zeitraum eines Jahres sind $\delta t \cdot p/100 \cdot y_0$. Damit ergibt sich insgesamt:

$$y_1 \quad = \quad y_0 + \delta t \cdot p/100 \cdot y_0 \ ,$$

und die Funktion f der rechten Seite ist $f(y) = p/100 \cdot y$.

c) Die analytische Lösung des AWPs lautet $y(t) = y_0 e^{\frac{pt}{100}}$. Das erkennt man entweder schon direkt aus dem AWP oder rechnet es beispielsweise mit der Separation der Variablen nach ($\dot{y}(t) = p/100 \cdot y(t) \ = \ F(t) \cdot G(y)$ mit $F(t) = p/100$ und $G(y) = y$, sowie t_0=0):

$$\begin{aligned}
\frac{dy}{dt} &= F(t) \cdot G(y) \\
\frac{dy}{G(y)} &= F(t)dt \\
\int_{y_0}^{y} \frac{1}{G(\eta)} d\eta &= \int_{t_0}^{t} F(\tau) d\tau \\
\int_{y_0}^{y} \frac{1}{\eta} d\eta &= \int_{t_0}^{t} \frac{p}{100}\, d\tau \\
ln(y) - ln(y_0) &= \frac{p}{100} \cdot t \\
y(t) &= y_0 \cdot e^{\frac{pt}{100}}, \qquad t \in [0; b].
\end{aligned}$$

d) Wir verwenden nun 4 Schritte des Euler-Verfahrens zur kleineren Schrittweite $\delta t = 1/4$ (Jahr). Die Berechnungsvorschrift für y_{k+1} lautet dann:

$$y_{k+1} \quad = \quad y_k + \delta t \cdot \frac{p}{100} \cdot y_k \quad = \quad \left(1 + \frac{\delta t\ p}{100}\right) \cdot y_k .$$

Wenn wir diese Definition rekursiv einsetzen, so erhalten wir die explizite Formel zur direkten Berechnung von y_{k+1} aus y_0:

$$y_{k+1} \quad = \quad \left(1 + \frac{\delta t\ p}{100}\right) \cdot y_k \quad = \quad \left(1 + \frac{\delta t\ p}{100}\right)^2 \cdot y_{k-1} \quad = \quad \ldots \quad = \quad \left(1 + \frac{\delta t\ p}{100}\right)^{k+1} \cdot y_0 .$$

e) Für die konkreten Werte y_0=100000 Euro und $p = 2$ erhalten wir folgende Gesamtguthabenswerte y_{end} nach einem Jahr ($b = 1$):

$$\begin{aligned}
\text{Diga-Bank}: y_{\text{end}} &= 10^5 + 1 \cdot 2/100 \cdot 10^5 \ = \ 102000,00 \ , \\
\text{Spaßkasse}: \ y_{\text{end}} &= y_4 \ = \ \left(1 + \frac{\frac{1}{4} \cdot 2}{100}\right)^4 \cdot 10^5 \ = \ 102015,05 \ , \\
\text{Anal. Lsg.}: \ y_{\text{end}} &= 10^5 \cdot e^{\frac{2}{100}} \ = \ 102020,13 \ .
\end{aligned}$$

Man erkennt, dass mit ca. 15 Euro (0,015%) ein nicht unerheblicher Unterschied im Endkapital zwischen Diga und Spaßkasse vorliegt. Die Differenz der Spaßkasse zur analytischen Lösung beträgt nur etwa 5 Euro.

4) Instabilität der Mittelpunktsregel

Neben den bisher betrachteten Einschrittverfahren gibt es eine weitere Klasse der sogenannten Mehrschrittverfahren. Deren Grundidee ist, die in früheren als dem letzten Schritt bereits

berechneten Approximationen nicht wegzuwerfen sondern mitzunutzen.
Ein einfacher Vertreter der Mehrschrittverfahren ist die Mittelpunktsregel (MPR). Die MPR ist ein Zweischrittverfahren der folgenden Form:

$$y_{k+1} \quad = \quad y_{k-1} + 2\delta t \cdot f(t_k, y_k), \qquad k = 1, \ldots, N-1$$

Damit ein Mehrschrittverfahren konvergiert, ist zusätzlich zur Konsistenz nun auch eine Stabilitätsbedingung notwendig. Wir wollen in dieser Aufgabe an der MPR beobachten, was passiert, wenn diese Stabilität verletzt wird.

Dazu betrachten wir das AWP

$$\begin{aligned} \dot{y}(t) &= -y(t) \\ y(0) &= 1. \end{aligned} \tag{6}$$

Zu Beginn des Verfahrens liegt nur der Anfangswert y_0 vor. Somit benötigen wir noch einen weiteren Wert y_1, bevor die Mittelpunktsregel gestartet werden kann. Typischerweise benutzt man zur Erzeugung dieses Wertes einfache Einschrittverfahren. Im Folgenden wollen wir daher y_1 mit Hilfe eines Schrittes des expliziten Euler-Verfahrens berechnen.

a) Berechnen Sie die analytische Lösung des AWPs (6)!

b) Wenden Sie die Mittelpunktsregel für $t \in [0; 6]$, $N = 3$ und $\delta t = 2$ auf das AWP (6) an! Was stellen Sie im Vergleich mit der analytischen Lösung fest?

c) Berechnen Sie nun die Ergebnisse mit halber Schrittweite ($N = 6$ und $\delta t = 1$) und vergleichen Sie sie wieder mit der analytischen Lösung!

Lösung:

a) Die analytische Lösung des AWPs lautet $y(t) = e^{-t}$. Das erkennt man entweder direkt aus dem AWP oder rechnet es beispielsweise mit der Separation der Variablen nach ($\dot{y}(t) = -y(t) \;=\; F(t) \cdot G(y)$ mit $F(t) = 1$ und $G(y) = -y$, sowie $y(t_0 = 0) = 1$):

$$\begin{aligned} \frac{dy}{dt} &= F(t) \cdot G(y) \\ \frac{dy}{G(y)} &= F(t)dt \\ \int_{y_0}^{y} \frac{1}{G(\eta)} d\eta &= \int_{t_0}^{t} F(\tau) d\tau \\ \int_{1}^{y} -\frac{1}{\eta} d\eta &= \int_{0}^{t} 1 \; d\tau \\ -ln(y) + ln(1) &= t \\ y(t) &= e^{-t}, \qquad t \in [0; b]. \end{aligned}$$

b) Für $N = 3$ und $\delta t = 2$ ergibt sich y_1 mit einem Startschritt des expliziten Euler-Verfahrens zu $y_1 = y_0 + \delta t \cdot f(t_0, y_0) \;=\; 1 + 2(-1) = -1$. Damit kann die Mittelpunktsregel als Zweischrittverfahren starten und ergibt:

$$\begin{aligned} y_2 &= y_0 + 2\delta t \cdot f(t_1, y_1) &&= 1 - 2 \cdot 2(-1) = 5, \\ y_3 &= y_1 + 2\delta t \cdot f(t_2, y_2) &&= -1 - 2 \cdot 2(5) = -21. \end{aligned}$$

Man erkennt, dass die Werte betragsmäßig größer werden, aber das Vorzeichen wechselt. Dagegen beschreibt die analytische Lösung ein gleichmäßiges Abklingen gegen Null. Offensichtlich ist die Stabilitätsbedingung bei der Mittelpunktsregel hier verletzt.

c) Ein zu b) analoges Ergebnis erzielt man auch mit halb so großer Schrittweite $\delta t = 1$. Der Euler-Startschritt liefert $y_1 = y_0 + 1 \cdot (-y_0) = 0$. Damit berechnet man über die Vorschrift der Mittelpunktsregel:

$$\begin{aligned} y_2 &= y_0 + 2\delta t \cdot f(t_1, y_1) &&= 1 - 2 \cdot 0 = 1 \ , \\ y_3 &= y_1 + 2\delta t \cdot f(t_2, y_2) &&= 0 - 2 \cdot 1 = -2 \ , \\ y_4 &= y_2 + 2\delta t \cdot f(t_3, y_3) &&= 1 - 2 \cdot (-2) = 5 \ , \\ y_5 &= y_3 + 2\delta t \cdot f(t_4, y_4) &&= -2 - 2 \cdot 5 = -12 \ , \\ y_6 &= y_4 + 2\delta t \cdot f(t_5, y_5) &&= 5 - 2 \cdot (-12) = 29 \ . \end{aligned}$$

Die Instabilität lässt sich also nicht durch Verringerung der der Schrittweite beheben, sondern verschiebt lediglich den Zeitpunkt des Anschwellens der Werte nach hinten. Die Betrachtung weiterer Verfeinerungslevel mit Hilfe des matlab-Programms *vis_midpoint.m* macht dies nochmals deutlich (vgl. Abb. 1).

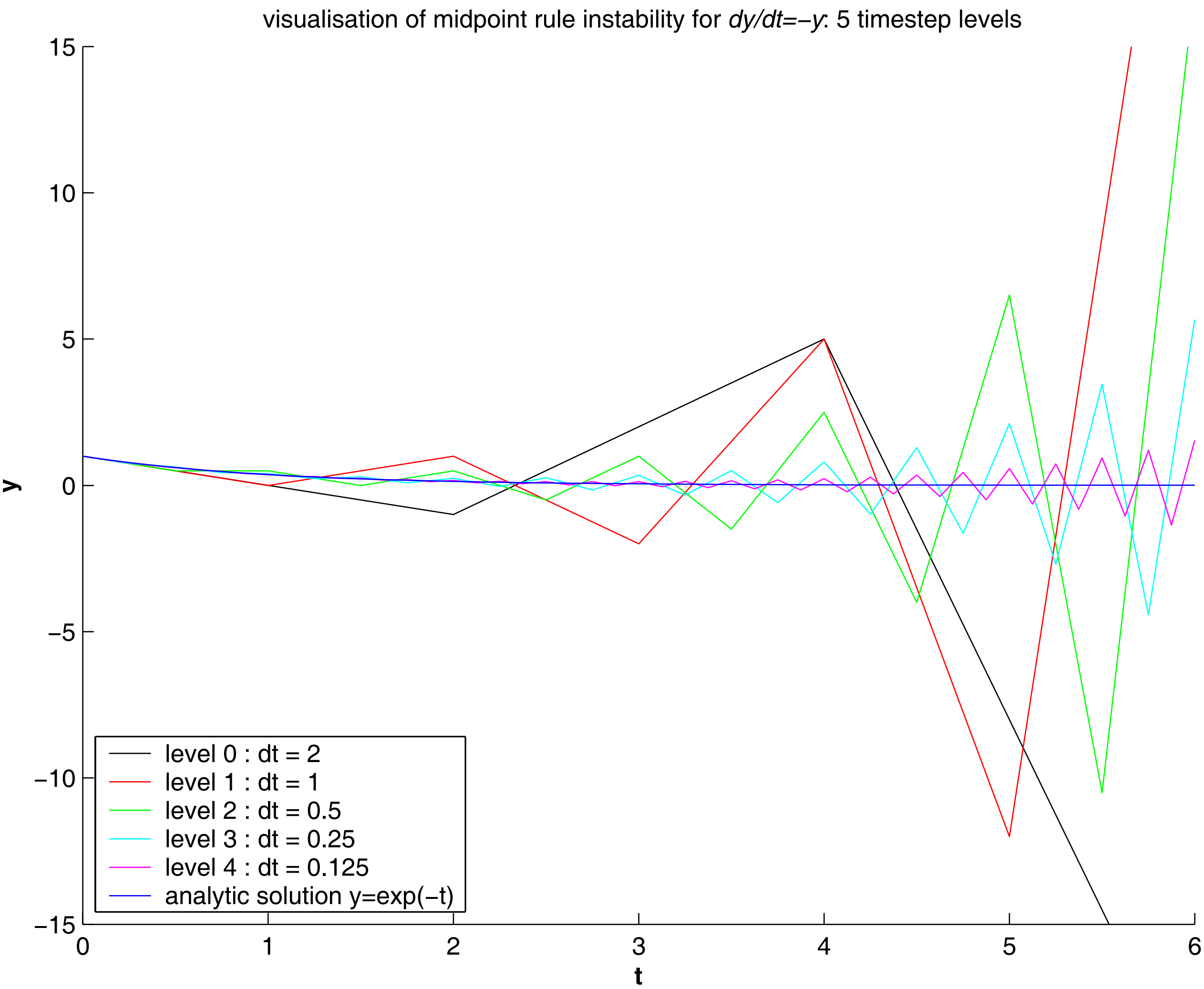

Abbildung 1: Visualisierung der Instabilität der Mittelpunktsregel: Fünf verschiedene Schrittweitenlevel und analytische Lösung.