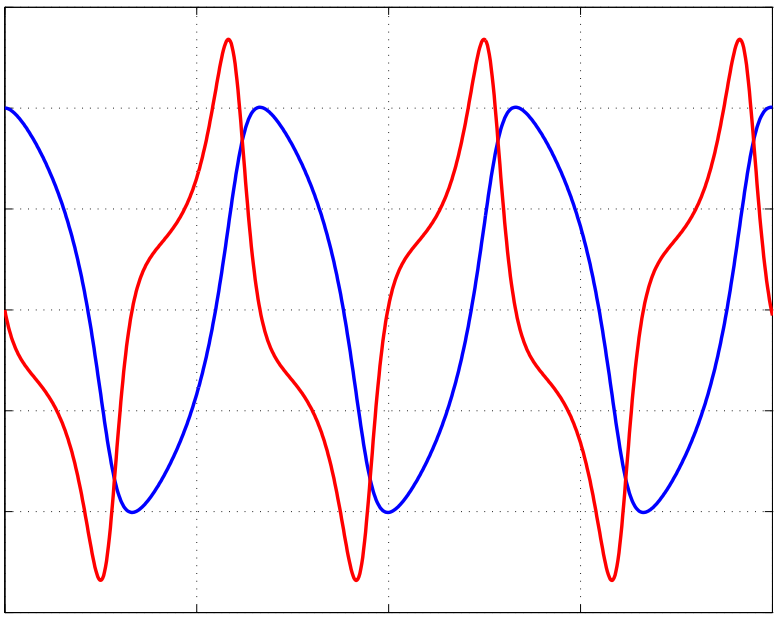
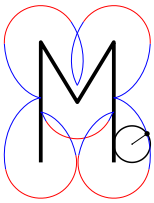
Mathematik–Online–Kurs

Numerische Methoden für Differentialgleichungen





http://www.mathematik-online.org/

Mathematik–Online–Kurs

Numerische Methoden für Differentialgleichungen

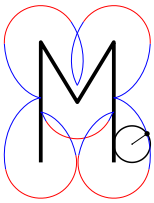
Stand: 25. Juli 2011

Konzipiert von K. Höllig iunter Mitwirkung von J. Hörner und M. Pfeil

© 2011 Mathematik-Online

Diese Veröffentlichung ist urheberrechtlich geschützt.

Weder Mathematik-Online noch einer der Autoren übernehmen Haftung für die Aktualität, Korrektheit, Vollständigkeit oder Qualität dieser Veröffentlichung. Haftungsansprüche, welche sich auf Schäden materieller oder ideeller Art beziehen, die durch die Nutzung oder Nichtnutzung der dargebotenen Informationen bzw. durch die Nutzung fehlerhafter und unvollständiger Informationen verursacht wurden, sind grundsätzlich ausgeschlossen.



http://www.mathematik-online.org/

Vorwort

Die Broschüre wurde im Rahmen von „Mathematik-Online“ aus dem entsprechenden Kurs-Modul erstellt. Sie richtet sich sowohl an Studierende der Mathematik als auch der Ingenieur- und Naturwissenschaften. Ergänzend findet man in der Rubrik „Aufgaben“ von Mathematik-Online auf die Inhalte des Kurs-Moduls abgestimmte Übungsaufgaben.

Sehr herzlich möchte ich Herrn Hörner und Frau Dr. Pfeil danken, die bei der Entwicklung dieses Kurses mitgewirkt haben. Die gemeinsame Arbeit an dem Projekt hat mir viel Freude bereitet, und ich wünsche den Lesern viel Spaß mit „Mathematik Online“ und Erfolg in ihrem Studium.

Stuttgart, im Januar 2008 Klaus Höllig

Inhaltsverzeichnis

Kapitel 1

Einschrittverfahren

1.1 Numerische Lösung von Differentialgleichungssystemen

Bei der numerischen Berechnung der Lösung $(u_1(t), \ldots, u_d(t))^{\mathrm{t}}$ eines Anfangswertproblems

$$u' = f(t,u)\,, \quad u(t_0) = u_0\,,$$

für ein System von Differentialgleichungen werden sukzessive Näherungen

$$u_\ell^h \approx u(t_\ell^h), \quad t_{\ell+1}^h = t_\ell^h + h_\ell\,,$$

für eine Folge $h_1, h_2, \ldots$ hinreichend kleiner Schrittweiten berechnet.
Ein Einschrittverfahren hat die Form

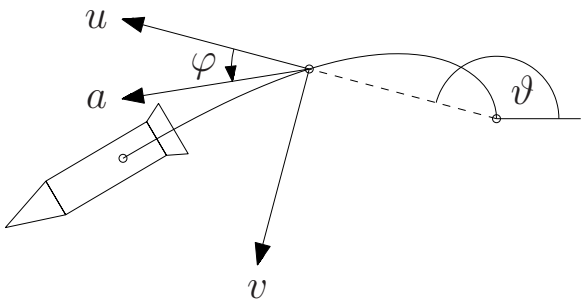
$$u_{\ell+1}^h = u_\ell^h + h_\ell \Phi\left(t_\ell^h, h_\ell, u_{\ell+1}^h, u_\ell^h, f\right)\,,$$

d.h. nur die letzte der bereits bestimmten Näherungen wird zur Berechnung von $u_{\ell+1}^h$ herangezogen. Bei einem n-Schrittverfahren hängt die Verfahrensfunktion Φ von $u_{\ell-n+1}^h, \ldots, u_\ell^h, u_{\ell+1}^h$, also den n letzten Näherungen ab. Hängt Φ nicht von $u_{\ell+1}^h$ ab, so bezeichnet man das Verfahren als explizit, sonst als implizit.
Bei der Wahl eines Verfahrens müssen Aufwand und Genauigkeit gegeneinander abgewogen werden. Darüber hinaus ist eine Schrittweitensteuerung für die Effizienz der Berechnung von entscheidender Bedeutung.

Beispiel:

In Polarkoordinaten wird die Flugbahn eines Raumschiffs durch die Differentialgleichungen

$$\begin{aligned} r' &= u \\ \vartheta' &= \frac{v}{r} \\ u' &= \frac{v^2}{r} + \frac{\gamma_E}{r^2} + a\cos\varphi \\ v' &= -\frac{uv}{r} + a\sin\varphi \end{aligned}$$



beschrieben, wobei $\gamma_E = 3.9869 \cdot 10^{14} \frac{\mathrm{N\,m^2}}{\mathrm{kg}}$ die Gravitationskonstante der Erde, (r, ϑ) die Position in Polarkoordinaten, (u, v) die radialen und tangentialen Geschwindigkeitskomponenten und (a, φ) der Betrag und Richtung der Beschleunigung des Antriebssystems bezogen auf den Ortsvektor sind.

1.2 Euler-Verfahren

Ein elementares Verfahren zur Approximation der Lösung $(u_1(t), \ldots, u_d(t)^{\mathrm{t}}$ einer Differentialgleichung $u' = f(t, u)$ basiert auf der Taylor-Approximation

$$u(t+h) = u(t) + h \underbrace{u'(t)}_{f(t,u(t))} + O(h^2)\,.$$

Bei Vernachlässigung von Termen zweiter Ordnung erhält man das Verfahren

$$\begin{aligned} u^h_{\ell+1} &= u^h_\ell + h_\ell f(t^h_\ell, u^h_\ell) \\ u^h_0 &= u(t_0)\,. \end{aligned}$$

Für stetig differenzierbare Funktionen f gilt für den Fehler bei konstanter Schrittweite

$$|u(t^h_\ell) - u^h_\ell| = O(h)$$

auf einem festen Intervall $[t_0, t_0 + T] \ni t^h_\ell = t_0 + \ell h$.

Beweis:

Zunächst wird der Diskretisierungsfehler abgeschätzt, d.h. die Diskrepanz,

$$\Delta_\ell = \frac{u(t^h_{\ell+1}) - u(t^h_\ell)}{h} - f(t^h_\ell, u(t_\ell))\,,$$

die beim Einsetzen der exakten Lösung in das Verfahren entsteht. Wegen $f(t^h_\ell, u(t^h_\ell)) = u'(t^h_\ell)$ gilt nach der Formel für das Taylor-Restglied

$$|\Delta_\ell| \le \frac{1}{2} h d$$

mit d einer Schranke für $|u''(t)|\,,\ t \in [t_0, t_0 + T]$.
Als nächstes muss die Auswirkung einer Störung im Argument von f berücksichtigt werden:

$$|f(t, u) - f(t, v)| \le c|u - v|$$

mit c einer Schranke für $|f_u|$ in einer Umgebung der exakten Lösung ($|u - v| \le \delta$).
Mit Hilfe der beiden obigen Ungleichungen kann nun eine rekursive Abschätzung für den Fehler hergeleitet werden. Durch Subtraktion der Gleichungen

$$\begin{aligned} u(t^h_{\ell+1}) &= u(t^h_\ell) + h f(t^h_\ell, u(t^h_\ell)) + h\Delta_\ell \\ u^h_{\ell+1} &= u^h_\ell + h f(t^h_\ell, u^h_\ell) \end{aligned}$$

folgt mit $e_\ell = |u(t^h_\ell) - u^h_\ell|$

$$e_{\ell+1} \le e_\ell + c h e_\ell + \frac{d}{2} h^2\,.$$

Durch Iteration dieser Ungleichung erhält man

$$\begin{aligned} e_\ell &\le (1+ch)e_{\ell-1} + \frac{d}{2}h^2 \\ &\le (1+ch)^2 e_{\ell-2} + (1+ch)\frac{d}{2}h^2 + \frac{d}{2}h^2 \\ &\le \dots \\ &\le (1+ch)^\ell e_0 + \frac{d}{2}h^2(1+(1+ch)+\dots+(1+ch)^{\ell-1})\,. \end{aligned}$$

Wegen $e_0 = 0$ und $1+ch \le e^{ch}$ ist also

$$e_\ell \le \frac{d}{2}\,h^2\,\frac{(1+ch)^\ell - 1}{ch} \le \left(\frac{d}{2c}e^{ch\ell}\right)h\,,$$

wobei der Vorfaktor $\frac{d}{2c}e^{ch\ell}$ wegen $h\ell \le T$ unabhängig von h beschränkt ist.

Beispiel:

Für das Modellproblem

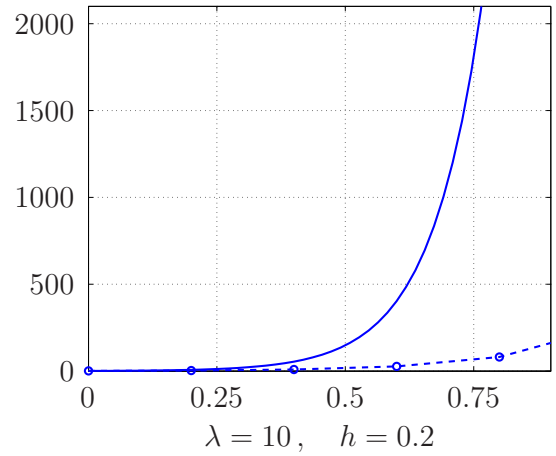
$$u' = \lambda u\,, \quad u(0) = 1$$

hat das Euler-Verfahren mit konstanter Schrittweite h die Form

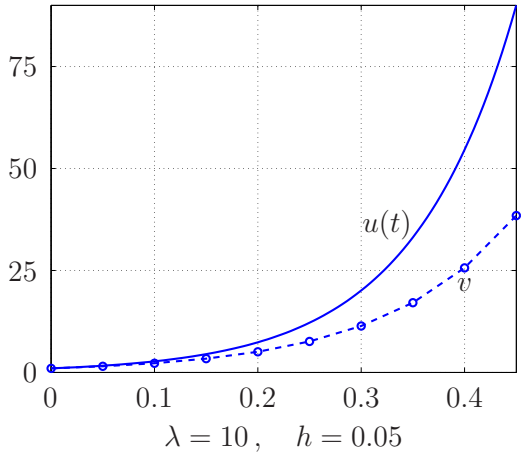
$$u^h_{\ell+1} = u^h_\ell + h\lambda u^h_\ell\,,$$

d.h.

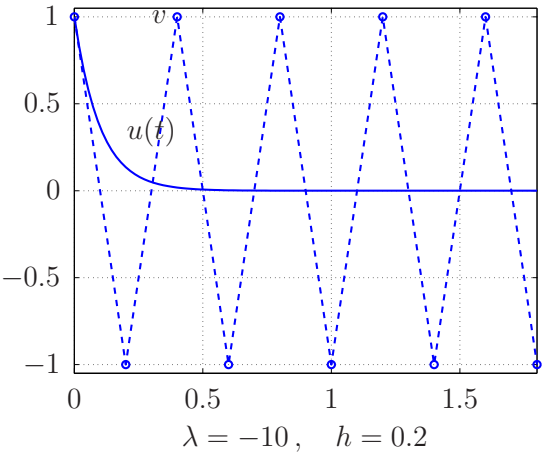
$$u^h_\ell = (1+h\lambda)^\ell\,.$$



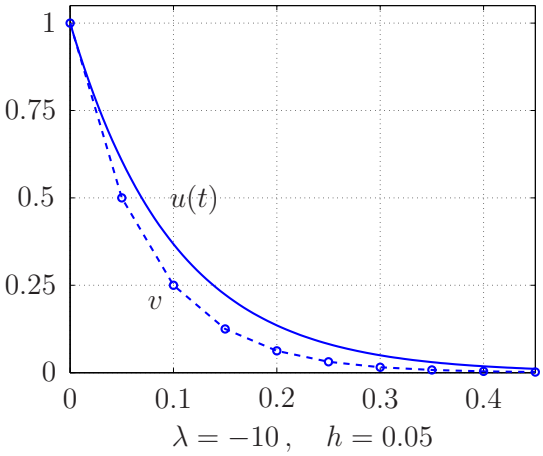
$\lambda = 10\,, \quad h = 0.2$



$\lambda = 10\,, \quad h = 0.05$



$\lambda = -10\,, \quad h = 0.2$



$\lambda = -10\,, \quad h = 0.05$

Die Abbildungen veranschaulichen zwei typische Probleme bei der Approximation gewöhnlicher Differentialgleichungen. Für großes positives λ wächst der Fehler exponentiell mit der Länge des Zeitintervalls. Mit $t = \ell h$ ist

$$u(t) - u_\ell^h = e^{\lambda t} - (1 + h\lambda)^{t/h} .$$

Schreibt man $1 + h\lambda = e^{\ln(1+h\lambda)}$, so erhält man mit dem Mittelwertsatz für die rechte Seite

$$\left(\lambda t - \frac{t}{h}\ln(1 + h\lambda)\right) e^{\tau} , \quad \frac{t}{h}\ln(1 + h\lambda) \le \tau \le \lambda t ,$$

also nach Taylor-Entwicklung

$$\left(\frac{1}{2}\lambda^2 th + O\left(h^2\right)\right) e^{\lambda t + O(h)} \approx h\left(\frac{1}{2}\lambda^2 t e^{\lambda t}\right) .$$

Ein derartiges Wachstum ist typisch für Differentialgleichungssysteme mit großen Realteilen der Eigenwerte der Jacobi-Matrix f_u.
Für großes negatives λ kommt es zu exponentiell anwachsenden Oszillationen der numerischen Lösung, wenn die Schrittweite h nicht klein genug gewählt wird, also einem völlig anderen qualitativen Verhalten als bei der exakten Lösung, die exponentiell abfällt. Für $\lambda = -10$ ist $h < 1/5$ notwendig, damit $\left|u_\ell^h\right| \to 0$. Für allgemeine Differentialgleichungssysteme ist die numerische Approximation besonders dann kritisch, wenn die Jacobi-Matrix f_u Eigenwerte mit großem negativen Realteil hat (steife Systeme).

1.3 Diskretisierungsfehler

Sei

$$u(t) \approx v \to w = v + h\Phi(t, h, w, v, f) \approx u(t + h)$$

ein Schritt eines Einschrittverfahrens zur Approximation der Lösung $(u_1(t), \ldots, u_d(t))^t$ eines Differentialgleichungssystems

$$u' = f(t, u) .$$

Dann bezeichnet man

$$\Delta(t, h) = \frac{u(t + h) - u(t)}{h} - \Phi(r, h, u(t + h), u(t), f)$$

als Diskretisierungsfehler. Der Vektor Δ ist also die Diskrepanz, die beim Einsetzen der exakten Lösung in die Verfahrensfunktion Φ entsteht.
Gilt bei glattem f

$$\|\Delta(t, h)\| = O(h^m)$$

für alle glatten Lösungen u, und ist der Exponent m bestmöglich, so hat das Verfahren die Ordnung m. Ein Verfahren mit einer Ordnung $m \ge 1$ bezeichnet man als konsistent.
Die Ordnung kann mit Taylor-Entwicklung bestimmt werden. Sie ist die kleinste natürliche Zahl m, für die die Ableitung

$$\partial_h^m \Delta(t, h)_{|h=0} = \frac{1}{m + 1} u^{(m+1)}(t) - \partial_h^m \Phi(t, h, u(t + h), u(t), f)_{|h=0}$$

nicht verschwindet. Dabei können Ableitungen von u mit Hilfe des Differentialgleichungssystems eliminiert werden. Benutzt man die Abkürzungen

$$f_{i,t,\ldots,t}^{j,k,\ldots} = \partial_t \ldots \partial_t \frac{\partial}{\partial u_j} \frac{\partial}{\partial u_k} \cdots f_i(t, u(t))$$

so gilt nach der Kettenregel

$$\begin{aligned} u_i'(t) &= f_i \\ u_i''(t) &= f_{i,t} + \sum_{j=1}^{d} f_i^j f_j \\ u_i'''(t) &= f_{i,t,t} + \sum_{j=1}^{d} (2 f_{i,t}^j f_j + f_i^j f_{j,t}) + \sum_{j,k=1}^{d} (f_i^{j,k} f_j f_k + f_i^j f_j^k f_k) \end{aligned}$$

usw..

Beweis:

Durch Taylor-Entwicklung erhält man für den ersten Term des Diskretisierungsfehlers

$$\frac{u(t+h) - u(t)}{h} = u'(t) + u''(t)\frac{h}{2} + u'''(t)\frac{h^2}{6} + \cdots .$$

Damit folgt die Formel für $\partial_h^m \Delta(t,h)_{|h=0}$.
Um die Formeln für die Ableitungen der Lösung u herzuleiten, werden die Komponenten des Differentialgleichungssystems

$$u_i'(t) = f_i(t, u_1(t), \ldots, u_d(t))$$

differenziert. Nach der Kettenregel ist

$$u_i'' = \partial_t f_i + \sum_{j=1}^{d} \left(\frac{\partial}{\partial u_j} f_i \right) u_i' = f_{i,t} + \sum_j f_i^j f_j$$

wegen $u_j' = f_j$. Jede Funktion g auf der rechten Seite hängt wiederum von den Variablen $t, u_1(t), \ldots, u_d(t)$ ab, kann also in derselben Weise abgeleitet werden:

$$\frac{d}{dt} g = g_t + \sum_k g^k f_k \, .$$

Mit der Kettenregel folgt so die Formel für u_i'', und analog erhält man alle weiteren Ableitungen.

Beispiel:

Als Beispiel wird das durch

$$u(t) \approx v \to w = v + h \underbrace{f(t + h/2, v + (h/2) f(t,v))}_{\Phi(t,h,v,f)} \approx u(t+h)$$

definierte Verfahren betrachtet. Es basiert auf einer symmetrischen Approximation der Ableitung an der Stelle $t + h/2$:

$$\frac{u(t+h) - u(t)}{h} \approx u'(t + h/2) = f(t + h/2, u(t + h/2)) .$$

Dabei wird $u(t + h/2)$ durch die lineare Taylor-Approximation ersetzt:

$$u(t + h/2) \approx u(t) + \frac{h}{2}u'(t)$$

mit $u'(t) = f(t, u(t))$.
Einsetzen der exakten Werte für v und w ergibt den Diskretisierungsfehler

$$\Delta(t, h) = \frac{u(t+h) - u(t)}{h} - f(t + h/2, u(t) + (h/2)f(t, u(t))) .$$

Zur Bestimmung der Ordnung werden nun sukzessive die Ableitungen bzgl. h gebildet.

$\underline{m = 0}$:

$$\Delta(t, 0) = u'(t) - f(t, u(t)) = 0 .$$

$\underline{m = 1}$: Mit $p = t + h/2$ und $q = u(t) + (h/2)f(t, u(t))$ ist

$$\partial_h \Phi_i = \partial_p f_i(p, q)p_h + \sum_{j=1}^{d} \frac{\partial}{\partial q_j} f_i(p, q)\partial_h q_h .$$

Wegen $p_h = 1/2$ und $\partial_h q_j = \frac{1}{2} f_j(p, q)$ erhält man

$$\partial_h \Delta_i(t, h)_{|h=0} = \frac{1}{2}u_i''(t) - \frac{1}{2}f_{i,t} + \frac{1}{2}\sum_{j=1}^{d} f_i^j f_j = 0 .$$

$\underline{m = 2}$: Nochmaliges Ableiten ergibt

$$\partial_h^2 \Phi_i = \partial_p^2 f_i(p, q)p_h^2 + r(t, h) ,$$

wobei r keine zweiten partiellen Ableitungen von f nach p enthält. Damit kann

$$\partial_h^2 \Delta_i(t, h)_{|h=0} = \frac{1}{3}u_i'''(t) - \frac{1}{4}f_{i,t,t} - r(t, 0)$$

nicht Null sein, denn $u'''(t) = f_{i,t,t} + \cdots$. Das Verfahren hat also die Ordnung $m = 2$.
Dass die Ordnung 2 bestmöglich ist, kann man auch anhand des Modellproblems

$$u' = u$$

verifizieren. Das Verfahren hat in diesem speziellen Fall die Form

$$w = v + h(v + (h/2)v) = [1 + h + h^2/2]v .$$

Der Faktor $[\ldots]$ approximiert e^h nur mit der Ordnung $m = 2$.

1.4 Trapez-Regel

Integriert man die Lösung $(u_1(t), \ldots, u_d(t))^t$ eines Differentialgleichungssystems

$$u' = f(t, u)\,,$$

so folgt

$$u(t+h) = u(t) + \int_t^{t+h} f(s, u(s))\, ds\,.$$

Das Integral kann durch die Trapezregel angenähert werden. Ein Schritt $u(t) \approx v \to w \approx u(t+h)$ des resultierenden Verfahrens hat dann die Form

$$w = v + \frac{h}{2}(f(t, v) + f(t+h, w))\,.$$

Die Trapezregel ist ein implizites Einschrittverfahren der Ordnung 2.
Zur Durchführung eines Verfahrensschritts $v \to w$ muss im allgemeinen ein nichtlineares Gleichungssystem gelöst werden. Näherungsweise kann dies durch einige wenige Schritte einer Fixpunktiteration

$$w^{\text{neu}} = v + \frac{h}{2}(f(t, v) + f(t+h, w^{\text{alt}}))$$

mit Startwert $w = v$ geschehen.
Für ein lineares Differentialgleichungssystem

$$u' = A(t)u + b(t)$$

ist die Trapezregel einfacher realisierbar. Ein Schritt erfordert lediglich die Lösung des linearen Gleichungssystems

$$\left(E - \frac{h}{2}A(t+h)\right) w = \left(E + \frac{h}{2}A(t)\right) v + \frac{h}{2}(b(t) + b(t+h))$$

mit E der $d \times d$-Einheitsmatrix.

Beweis:

Der Diskretisierungsfehler der Trapezregel ist

$$\Delta(t, h) = \frac{u(t+h) - u(t)}{h} - \frac{1}{2}\left(f(t, u(t)) + f(t+h, u(t+h))\right).$$

Setzt man die Taylor-Entwicklungen

$$\begin{aligned} u(t+h) &= u(t) + u'(t)h + \frac{1}{2}u''(t)h^2 + \frac{1}{6}u'''(t)h^3 + O(h^4) \\ f(t+h, u(t+h)) = u'(t+h) &= u'(t) + u''(t)h + \frac{1}{2}u'''(t)h^2 + O(h^3) \end{aligned}$$

ein, so folgt

$$\Delta(t, h) = -\frac{1}{12}u'''(t)h^2 + O(h^3)\,.$$

Die Trapezregel hat also die Ordnung 2.

Beispiel:

Als Beispiel wird die Trapezregel auf das lineare Differentialgleichungssystem

$$u' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} u, \quad u(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

mit der Lösung $u(t) = (\cos t, \sin t)^t$ angewandt. Ein Schritt $u(t) \approx v \to w \approx u(t+h)$ hat die Form

$$w = v + \frac{h}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} (v + w)$$

bzw., aufgelöst nach w,

$$w = \underbrace{\begin{pmatrix} 1 & h/2 \\ -h/2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -h/2 \\ h/2 & 1 \end{pmatrix}}_{A(h)} v\,.$$

Man erkennt, dass $A(h)$ eine Drehmatrix ist:

$$A(h) = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}, \quad \tan(\vartheta/2) = h/2\,.$$

Folglich bleibt die euklidische Norm der numerischen Lösung bei einem Zeitschritt unverändert:

$$\|w\|_2 = \|v\|_2\,.$$

Das qualitative Verhalten der exakten Lösung u wird also durch die Trapezregel akurat modelliert. Bezeichnet $u^h \approx u(2\pi)$ die numerische Lösung nach $\ell = 2\pi/h$ Schritten, so ist

$$\arg u^h = \vartheta \cdot 2\pi/h$$

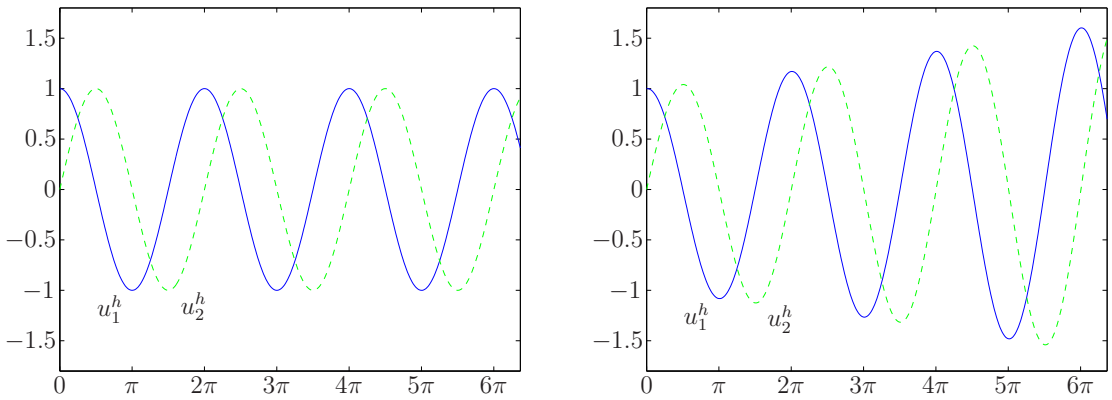
mit $\arg(x,y)^t$ dem Winkel in der Polardarstellung eines Vektors $(x,y)^t$. Wegen

$$\vartheta = 2\arctan(h/2) = h + O(h^3)$$

ist also

$$\arg u^h = 2\pi + O(h^3)\,.$$

Die Approximation ist genauer als erwartet.



Die Abbildung vergleicht die Trapezregel (links) mit der Approximation durch Eulers Verfahren (rechts) bei einer Schrittweite $h = 1/20$. Die Vorteile des impliziten Verfahrens sind offensichtlich. Insbesondere hat die Trapezregel ein besseres Langzeitverhalten.

1.5 Konvergenz von Einschrittverfahren

Für ein konsistentes Einschrittverfahren zur Lösung des Anfangswertproblems

$$u' = f(t,u),\ t_0 \le t \le t_0 + T, \quad u(t_0) = u_0\,,$$

lässt sich der Fehler durch den Diskretisierungsfehler abschätzen:

$$\|u_\ell^h - u(t_\ell^h)\| \le \text{const}\,\delta$$

für hinreichend kleine Schrittweiten h_ℓ. Dabei ist $t_{\ell+1}^h = t_\ell^h + h_\ell$, u_ℓ^h die numerische Approximation von $u(t_\ell^h)$ und

$$\delta = \max_{t_0 \le t_\ell^h < t_0+T} \|\Delta(t_\ell^h, h_\ell)\|$$

das Maximum des Diskretisierungsfehlers. Die Konstante const hängt von dem Differentialgleichungssystem, dem Verfahren und der Intervalllänge T, aber nicht von der Schrittweitenfolge h ab. Insbesondere gilt für ein Verfahren der Ordnung m:

$$\|u_\ell^h - u(t_\ell^h)\| = O(|h|^m)$$

mit $|h| = \max_{t_0 \le t_\ell^h < t_0+T} h_\ell$.

Beweis:

Ein Einschrittverfahren hat die Form

$$u_{\ell+1}^h = u_\ell^h + h\Phi(t_\ell^h, h, u_{\ell+1}^h, u_\ell^h, f)\,.$$

Dabei ist die Verfahrensfunktion $\Phi(t,h,w,v,f)$ und ihre Ableitungen in einer Umgebung der exakten Lösung u beschränkt:

$$\|\Phi\|, \left\|\frac{\partial\Phi}{\partial v}\right\|, \left\|\frac{\partial\Phi}{\partial v}\right\| \le c$$

für

$$\|u_\ell^h - u(t_\ell^h)\| \le \varepsilon, \quad t_0 \le t_\ell^h < t_0 + T,\ 0 \le h_\ell \le \varepsilon\,,$$

mit $\varepsilon > 0$.
Für den Fehler $e_\ell = \|u_\ell^h - u(t_\ell^h)\|$ wird nun die Abschätzung

$$e_\ell \le C\delta\left(e^{C(t_\ell^h - t_0)} - 1\right)$$

mit $C = \max(2, 4c)$ gezeigt. Die maximale Schrittweite h_ε wird so klein gewählt, dass die rechte Seite der Fehlerabschätzung für alle relevanten ℓ kleiner als $\varepsilon/2$ ist und

$$\max h_\ell \le \min(1/(2c), \varepsilon/(2c))\,.$$

Mit diesen Annahmen kann die Fehlerabschätzung nun induktiv gezeigt werden.
Da $u_0^h = u_0 = u(t_0)$, ist die Ungleichung für $\ell = 0$ trivialerweise erfüllt. Für den Induktionsschritt $\ell \to \ell + 1$ zeigt man zunächst, dass das nichtlineare Gleichungssystem

$$w = g(w) = u_\ell^h + h\Phi(t_\ell^h, h_\ell, w, u_\ell^h, f)$$

eine eindeutige Lösung $w = u^h_{\ell+1}$ besitzt. Dies folgt aus dem Banachschen Fixpunktsatz. Wegen

$$\|u^h_\ell - u(t^h_\ell)\| \le \varepsilon/2, \quad h_\ell \|\Phi\| \le \varepsilon/2$$

bildet g die Menge

$$D: \ \|w - u(t^h_\ell)\| \le \varepsilon$$

in sich ab. Weiter ist

$$\left\| \frac{\partial g}{\partial w} \right\| = h_\ell \left\| \frac{\partial \Phi}{\partial w} \right\| \le \frac{1}{2},$$

so dass die Kontraktionsbedingung ebenfalls erfüllt ist.
Zum Beweis der Fehlerabschätzung schreibt man den Diskretisierungsfehler in der Form

$$u(t^h_{\ell+1}) = u(t^h_\ell) + h_\ell \Phi(t^h_\ell, h_\ell, u(t^h_{\ell+1}), u(t^h_\ell), f) + h_\ell \Delta(t^h_\ell, h_\ell)$$

und zieht die Verfahrensgleichung

$$u^h_{\ell+1} = u^h_\ell + h_\ell \Phi(t^h_\ell, h_\ell, u^h_{\ell+1}, u^h_\ell, f)$$

ab. Benutzt man die Ungleichung

$$\|\varphi(\xi,\eta) - \varphi(\tilde\xi,\tilde\eta)\| \le \left(\max \left\| \frac{\partial \varphi}{\partial \xi} \right\| \right) \|\xi - \tilde\xi\| + \left(\max \left\| \frac{\partial \varphi}{\partial \eta} \right\| \right) \|\eta - \tilde\eta\|,$$

so gewinnt man daraus die Abschätzung

$$e_{\ell+1} \le e_\ell + c h_\ell (e_{\ell+1} + e_\ell) + h_\ell \delta .$$

Wegen $ch_\ell \le 1/2$ lässt sich diese Ungleichung nach $e_{\ell+1}$ auflösen, und man erhält

$$e_{\ell+1} \le \frac{1 + ch_\ell}{1 - ch_\ell} e_\ell + 2h_\ell \delta .$$

Aufgrund der Bedingung an $h\ell$ und der Induktionsvoraussetzung kann die rechte Seite nun durch

$$(1 + Ch_\ell)e_\ell + Ch_\ell\delta \le e^{Ch_\ell} C\delta(e^{C(t^h_\ell - t:0)} - 1) \ + \ Ch_\ell\delta = C\delta e^{C(t^h_{\ell+1} - t_0)} \ - \ C\delta e^{Ch_\ell} \ + \ Ch_\ell\delta$$

abgeschätzt werden. Die letzten beiden Terme sind

$$\le -C\delta(1 + Ch_\ell) + Ch_\ell\delta \le -C\delta ,$$

so dass man die gewünschte Ungleichung für $e_{\ell+1}$ erhält.

1.6 Runge-Kutta-Verfahren

Bei einem Zeitschritt

$$v \approx u(t) \to u(t+h) \approx w$$

eines n-stufigen Runge-Kutta-Verfahrens zur Approximation des Differentielgleichungssystems

$$u' = f(t,u)$$

werden zunächst die n Hilfsgrößen

$$y_i = f\left(t + c_i h, v + h\sum_{j=1}^{n} a_{i,j} y_j\right), \quad i = 1, \ldots, n,$$

mit $c_i = \sum_{j=1}^{n} a_{i,j}$ berechnet und dann die gewichtete Summe

$$w = v + h\sum_{j=1}^{n} b_j y_j$$

gebildet.
Die Approximationsordnung hängt von der Wahl der Verfahrensparameter

$$R = \left(\begin{array}{c|c} A & c \\ \hline b^{\mathrm{t}} & \end{array}\right) = \left(\begin{array}{ccc|c} a_{1,1} & \cdots & a_{1,n} & c_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,n} & c_n \\ \hline b_1 & \cdots & b_n & \end{array}\right)$$

ab.
Ist

$$a_{i,j} = 0, \quad j \geq i,$$

so lassen sich die Hilfsgrößen y_i sukzessive berechnen, und das Verfahren ist explizit. Andernfalls müssen die Hilfsgrößen simultan durch Lösen eines nichtlinearen Gleichungssystems bestimmt werden. Solche impliziten Runge-Kutta-Verfahren sind deshalb in erster Linie für lineare Differentialgleichungen geeignet.

1.7 Klassisches Runge-Kutta-Verfahren

Das klassische Runge-Kutta-Verfahren ist ein explizites vierstufiges Verfahren vierter Ordnung. Ein Zeitschritt

$$u(t) \approx v \to w \approx u(t+h)$$

zur Approximation des Differentialgleichungssystems $u' = f(t,u)$ hat die Form

$$\begin{aligned} y_1 &= f(t, v) \\ y_2 &= f(t + h/2, v + h y_1/2) \\ y_3 &= f(t + h/2, v + h y_2/2) \\ y_4 &= f(t + h, v + h y_3) \\ w &= v + h(y_1/6 + y_2/3 + y_3/3 + y_4/6). \end{aligned}$$

Die zugehörige Parametermatrix ist

$$R = \left(\begin{array}{c|c} A & c \\ \hline b^{\mathrm{t}} & \end{array}\right) = \left(\begin{array}{cccc|c} & & & & 0 \\ \frac{1}{2} & & & & \frac{1}{2} \\ 0 & \frac{1}{2} & & & \frac{1}{2} \\ 0 & 0 & 1 & & 1 \\ \hline \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \end{array}\right).$$

Beispiel:

Zur Illustration der Konvergenzordnung wird das klassische Runge-Kutta-Verfahren auf das Modellproblem

$$u' = \lambda u$$

angewandt. Die Hilfsgrößen für einen Schritt sind dann

$$\begin{array}{lcll} y_1 & = & f(t, v) & = \lambda v \\ y_2 & = & f(t + h/2, v + hy_1/2) & = \lambda(v + h\lambda v/2) = \lambda v + h\lambda^2 v/2 \\ y_3 & = & f(t + h/2, v + hy_2/2) & = \lambda v + h\lambda^2 v/2 + h^2\lambda^3 v/4 \\ y_4 & = & f(t + h, v + hy_3) & = \lambda v + h\lambda^2 v + h^2\lambda^3 v/2 + h^3\lambda^4 v/4 \end{array}$$

und man erhält

$$w = v + h\left(\frac{1}{6}y_1 + \frac{1}{3}y_2 + \frac{1}{3}y_3 + \frac{1}{6}y_4\right) = v(1 + h\lambda + h^2\lambda^2/2 + h^3\lambda^3/6 + h^4\lambda^4/24)\,.$$

Folglich ist

$$w = (\exp(h\lambda) + O((h\lambda)^5))v$$

in Übereinstimmung mit

$$u(t + h) = \exp(h\lambda)u(t)\,.$$

1.8 Gauß-Runge-Kutta-Verfahren

Das Gauß-Runge-Kutta-Verfahren der Ordnung $2n$ ist ein n-stufiges implizites Verfahren maximaler Ordnung $2n$. Seine Parameter

$$R = \left(\begin{array}{c|c} A & c \\ \hline b^{\mathrm{t}} & \end{array}\right)$$

werden mit Hilfe des Legendre-Polynoms p vom Grad n definiert. Die Stützstellen $c_1 < \cdots < c_n$ sind die auf das Intervall $[0, 1]$ transformierten Nullstellen von p, die Gewichte $b_1, \ldots, b_n$ die Integrale der Lagrange-Polynome,

$$b_j = \int_0^1 q_j(s)\,ds\,, \quad q_j(s) = \prod_{\substack{k=1 \\ k\neq j}}^{n} \frac{s - c_k}{c_j - c_k}\,.$$

und

$$a_{j,k} = \int_0^{c_j} q_k(s)\,ds\,.$$

Da für ein Verfahrensschritt ein nichtlineares Gleichungssystem gelöst werden muss, sind die Gauß-Runge-Kutta-Verfahren vor allem für lineare Differentialgleichungssysteme geeignet.

Beweis:

Die der Konstruktion des Verfahrens zugrunde liegende Idee lässt sich für eine skalare Differentialgleichung erläutern. Für die Realisierung eines Schritts $u(t) \approx v \to w \approx u(t+h)$ des Gauß-Runge-Kutta-Verfahrens wird ein Interpolationspolynom p konstruiert, dessen Ableitung $p'(hc_j)$ an den Stützstellen jeweils eine Approximation für $u'(t+hc_j)$ ist, d.h ein Polynom, das das Gleichungssystem

$$\begin{aligned} p(0) &= v \\ p'(hc_j) &= f(t+hc_j, p(hc_j))\,, \quad j=1,\ldots,n\,, \end{aligned}$$

erfüllt. Dann wird $w = p(h)$ gesetzt.
Die Ableitung des Polynoms hat die Lagrange-Darstellung

$$p'(s) = \sum_{k=1}^{n} p'(hc_k) q_k(s/h)\,,$$

und Integration in Verbindung mit der ersten Gleichung liefert

$$p(hc_j) = v + \int_0^{hc_j} \sum_{k=1}^{n} p'(hc_k) q_k(s/h)\, ds = v + \sum_{k=1}^{n} p'(hc_k) \int_0^{hc_j} q_k(s/h)\, ds\,.$$

Setzt man dies in die weiteren Gleichungen des Systems ein, so erhält man

$$p'(hc_j) = f\left(t + hc_j, v + \sum_{k=1}^{n} p'(hc_k) \int_0^{hc_j} q_k(s/h)\, ds \right)\,, \quad j=1,\ldots,n\,.$$

Dieses Gleichungssystem hat dieselbe Form wie die Definition der Hilfsgrößen y_i eines Runge-Kutta-Verfahrens mit

$$y_j = p'(hc_j)\,, \quad a_{j,k} = \frac{1}{h} \int_0^{hc_j} q_k(s/h)\, ds = \int_0^{c_j} q_k(s)\, ds\,.$$

Da die Lagrange-Polynome q_k zu 1 summieren, ist die Bedingung

$$\sum_{k=1}^{n} a_{j,k} = \sum_{k=1}^{n} \int_0^{c_j} q_k(s)\, ds = \int_0^{c_j} \sum_{k=1}^{n} q_k(s)\, ds = \int_0^{c_j} 1\, ds = c_j$$

erfüllt. Schließlich ist

$$w = p(h) = v + \int_0^h p'(s)\, ds = v + h \sum_{j=1}^{n} p'(hc_j) \int_0^1 q_j(s)\, ds\,,$$

in Übereinstimmung mit der Definition von b_j.
Definitionsgemäß gilt für den Diskretisierungsfehler der Gauß-Runge-Kutta-Verfahrens für eine glatte Lösung u

$$\begin{aligned} h\Delta(t,h) &= u(t+h) - u(t) - h \sum b_j y_j \\ &= \int_t^{t+h} f(s,u(s))\, ds - \int_0^h p'(s)\, ds\,. \end{aligned}$$

Da p' die Funktion f an den Gauß-Knoten interpoliert stimmt $\int_0^h p'$ mit der Gaußformel für $\int_t^{t+h} f$ überein. Die Differenz hat damit die Ordnung $O(h^{2n})$, wobei der zusätzliche Faktor h aus der Transformation des Intervalls $[0,1]$ auf $[0,h]$ resultiert.
Die Argumentation im vektorwertigen Fall ist vollkommen analog; die Polynominterpolation wird separat in jeder Komponente durchgeführt.

Beispiel:

Für das 2-stufige Gauß-Runge-Kutta-Verfahren werden zunächst die Nullstellen des Legendre-Polynoms vom Grad 2, $p_2(x) = 3x^2/2 - 1/2$, benötigt: $x = \pm\sqrt{3}/3$. Durch Transformation auf das Intervall $[0,1]$ erhält man

$$c_1 = \frac{1}{2} - \frac{1}{6}\sqrt{3}\,, \quad c_2 = \frac{1}{2} + \frac{1}{6}\sqrt{3}\,.$$

Die Lagrange-Polynome zu diesen Stützstellen sind

$$q_1(s) = \frac{s - c_2}{c_1 - c_2} = \frac{1}{2} + \frac{1}{2}\sqrt{3} - \sqrt{3}s\,, \quad q_2(s) = \frac{s - c_1}{c_2 - c_1} = \frac{1}{2} - \frac{1}{2}\sqrt{3} + \sqrt{3}s\,.$$

Die Gewichte sind symmetrisch und summieren zur Intervalllänge, also ist $b_1 = b_2 = 1/2$. Die Parameter $a_{j,k}$ ergeben sich schließlich durch Integration von q_k über das Intervall $[0, c_j]$. Beispielsweise ist

$$a_{1,2} = \int_0^{c_1} q_2(s)\,ds = \int_0^{\frac{1}{2}-\frac{1}{6}\sqrt{3}} \left(\frac{1}{2} - \frac{1}{2}\sqrt{3} + \sqrt{3}s\right) ds = \frac{1}{4} - \frac{1}{6}\sqrt{3}\,.$$

Insgesamt erhält man die Parametermatrix

$$R = \left(\begin{array}{c|c} A & c \\ \hline b^{\mathrm{t}} & \end{array}\right) = \left(\begin{array}{cc|c} \frac{1}{4} & \frac{1}{4} - \frac{1}{6}\sqrt{3} & \frac{1}{2} - \frac{1}{6}\sqrt{3} \\ \frac{1}{4} + \frac{1}{6}\sqrt{3} & \frac{1}{4} & \frac{1}{2} + \frac{1}{6}\sqrt{3} \\ \hline \frac{1}{2} & \frac{1}{2} & \end{array}\right).$$

Beispiel:

Für ein lineares Differentialgleichungssystem

$$u' = Q(t)u + p(t)$$

ist das Gleichungssystem für die Hilfsgrößen des Gauß-Runge-Kutta-Verfahrens linear. Mit $v \approx u(t)$ und $w \approx u(t+h)$ ist

$$y_j = Q(t + hc_j)\left(v + h\sum_{k=1}^{n} a_{j,k}y_k\right) + p(t + hc_j), \quad j = 1, \ldots, n\,,$$

und

$$w = v + h\sum_{j=1}^{n} b_j y_j\,.$$

Beispielsweise hat das lineare Gleichungssystem für ein 2-stufiges Verfahren die Blockform

$$\begin{pmatrix} E - ha_{1,1}Q_1 & -ha_{1,2}Q_1 \\ -ha_{2,1}Q_2 & E - ha_{2,2}Q_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} Q_1 v + p_1 \\ Q_2 v + p_2 \end{pmatrix}$$

mit $Q_j = Q(t + hc_j)$, $p_j = p(t + hc_j)$ und E der Einheitsmatrix der Dimension des Differentialgleichungssystems.

1.9 Eingebettetes Runge-Kutta-Verfahren

Ein Runge-Kutta-Verfahren mit Parametermatrix

$$R = \left(\begin{array}{c|c} A & c \\ \hline b^{\mathrm{t}} & \end{array}\right)$$

mit Ordnung m, das durch Ersetzen der Gewichte b durch $\tilde{b}$ zu einem Verfahren mit Ordnung $\tilde{m} \neq m$ wird, nennt man eingebettetes Verfahren.
Bei einem solchen Verfahren können ohne bedeutenden Mehraufwand zusätzliche Approximationen gewonnen werden, die zur Schrittweitensteuerung eingesetzt werden können.

Beispiel:

Ein häufig verwendetes eingebettetes Runge-Kutta-Verfahren ist das Verfahren von Dormand-Prince. Es hat die Parametermatrix

$$R = \left(\begin{array}{c|c} A & c \\ \hline b^{\mathrm{t}} & \end{array}\right) = \left(\begin{array}{ccccccc|c} & & & & & & & 0 \\ \frac{1}{5} & & & & & & & \frac{1}{5} \\ \frac{3}{40} & \frac{9}{40} & & & & & & \frac{3}{10} \\ \frac{44}{45} & -\frac{56}{15} & \frac{32}{9} & & & & & \frac{4}{5} \\ \frac{19372}{6561} & -\frac{25360}{2187} & \frac{64448}{6561} & -\frac{212}{729} & & & & \frac{8}{9} \\ \frac{9017}{3168} & -\frac{355}{33} & \frac{46732}{5247} & \frac{49}{176} & -\frac{5103}{18656} & & & 1 \\ \frac{35}{384} & 0 & \frac{500}{1113} & \frac{125}{192} & -\frac{2187}{6784} & \frac{11}{84} & & 1 \\ \hline \frac{35}{384} & 0 & \frac{500}{1113} & \frac{125}{192} & -\frac{2187}{6784} & \frac{11}{84} & 0 & \end{array}\right)$$

und die Ordnung 5. Ersetzt man den Vektor b^{t} durch

$$\tilde{b}^{\mathrm{t}} = \left(\frac{5179}{57600}, 0, \frac{7571}{16695}, \frac{393}{640}, -\frac{92097}{339200}, \frac{187}{2100}, \frac{1}{40}\right),$$

so ergibt sich ein Verfahren der Ordnung 4.

1.10 Ordnungsbedingungen für Runge-Kutta-Verfahren

Ein n-stufiges Runge-Kutta-Verfahren mit Parametermatrix

$$R = \left(\begin{array}{c|c} A & c \\ \hline b^{\mathrm{t}} & \end{array}\right),$$

hat die Ordnung m, wenn

$$c_j = \sum_k a_{j,k}$$

und die folgenden Bedingungen bis zur Ordnung m einschließlich erfüllt sind.

Ordnung 1:

$$\sum_j b_j = 1$$

Ordnung 2:

$$2\sum_{j,k} b_j a_{j,k} = 1$$

Ordnung 3:

$$\begin{aligned} 3\sum_{j,k,\ell} b_j a_{j,k} a_{j,\ell} &= 1 \\ 6\sum_{j,k,\ell} b_j a_{j,k} a_{k,\ell} &= 1 \end{aligned}$$

Ordnung 4:

$$\begin{aligned} 4\sum_{j,k,\ell,m} b_j a_{j,k} a_{j,\ell} a_{j,m} &= 1 \\ 8\sum_{j,k,\ell,m} b_j a_{j,k} a_{k,\ell} a_{j,m} &= 1 \\ 12\sum_{j,k,\ell,m} b_j a_{j,k} a_{k,\ell} a_{k,m} &= 1 \\ 24\sum_{j,k,\ell,m} b_j a_{j,k} a_{k,\ell} a_{\ell,m} &= 1 \end{aligned}$$

Beispiel:

Ein explizites 2-stufiges Runge-Kutta-Verfahren muss, um die Ordnung 2 zu haben, die Bedingungen

$$\begin{aligned} b_1 + b_2 &= 1 \\ b_2 a_{2,1} &= 1/2 \end{aligned}$$

erfüllen. Aufgrund der zweiten Gleichung müssen b_2 und $a_{2,1}$ von Null verschieden sein, und somit erhält man das einfachste Verfahren für $b_1 = 0$. Dies hat die Parametermatrix

$$R = \left(\begin{array}{c|c} A & c \\ \hline b^{\mathrm{t}} & \end{array}\right) = \left(\begin{array}{cc|c} & & 0 \\ 1/2 & & 1/2 \\ \hline 0 & 1 & \end{array}\right) .$$

Ein 2-stufiges explizites Verfahren kann nicht die Ordnung 3 haben, da hier stets $\sum b_j a_{j,k} a_{k,\ell} = 0 \neq 1/6$ gilt.
Ergänzt man das obige Verfahren um eine dritte Stufe, so müssen für Ordnung 3 folgende Gleichungen erfüllt sein:

$$\begin{aligned} \tilde{b}_1 + \tilde{b}_2 + \tilde{b}_3 &= 1\,, \\ \tilde{b}_2 + 2\,\tilde{b}_3\, c_3 &= 1\,, \\ 3\,\tilde{b}_2/4 + 3\,\tilde{b}_3\, c_3^2 &= 1\,, \\ 3\,\tilde{b}_3\, a_{3,2} &= 1\,. \end{aligned}$$

Wählt man nun $\tilde{b}_2 = 0$ ergeben sich aus der zweiten und dritten Gleichung $c_3 = 2/3$ und $\tilde{b}_3 = 3/4$. Damit folgt aus der vierten Gleichung $a_{3,2} = 4/9$ und damit $a_{3,1} = c_3 - a_{3,2} = 2/9$. Schließlich liefert die erste Gleichung noch $\tilde{b}_1 = 1/4$.
Es ergibt sich also das eingebettete Runge-Kutta-Verfahren

$$R = \left(\begin{array}{c|c} A & c \\ \hline b^{\mathrm{t}} & \\ \hline \tilde{b}^{\mathrm{t}} & \end{array}\right) = \left(\begin{array}{ccc|c} & & & 0 \\ 1/2 & & & 1/2 \\ 2/9 & 4/9 & & 2/3 \\ \hline 0 & 1 & 0 & \\ \hline 1/4 & 0 & 3/4 & \end{array}\right).$$

der Ordnung 2 bzw. 3 .

Kapitel 2

Mehrschrittverfahren

2.1 Lineares Mehrschrittverfahren

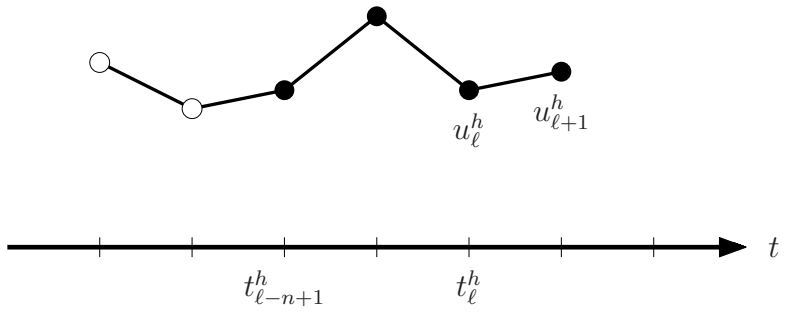
Ein lineares n-Schrittverfahren mit Parametern a_k und b_k zur Approximation der Lösung $(u_1(t), \ldots, u_d(t))^t$ eines Differentialgleichungssystems

$$u' = f(t, u),$$

hat die Form

$$u^h_{\ell+1} = \sum_{k=0}^{n-1} a_k u^h_{\ell-k} + h \sum_{k=-1}^{n-1} b_k f(t^h_{\ell-k}, u^h_{\ell-k})$$

mit $t^h_\ell = t_0 + \ell h$ und u^h_ℓ der Approximation von $u(t^h_\ell)$.



Wie in der Abbildung illustriert ist, basiert ein Schritt des Verfahrens auf den n zuletzt berechneten Approximationen. Zum Starten eines Mehrschrittverfahrens ist deshalb eine zusätzliche Prozedur erforderlich. Die ersten n Approximationen $u^h_0, \ldots, u^h_{n-1}$ können beispielsweise durch Taylor-Entwicklung oder mit Hilfe eines Einschrittverfahrens ausgehend von dem Anfangswert $u(t_0)$ berechnet werden.

Man unterscheidet zwischen expliziten und impliziten Mehrschrittverfahren, jenachdem ob der Koeffizient b_{-1} von $u^h_{\ell+1}$ Null oder ungleich Null ist. Explizite Mehrschrittverfahren benötigen nur eine Auswertung der Funktion f pro Schritt. Sie sind deshalb sehr effizient. Implizite Verfahren sind zwar etwas aufwändiger zur implementieren, haben jedoch im allgemeinen bessere Stabilitätseigenschaften.

Beispiel:

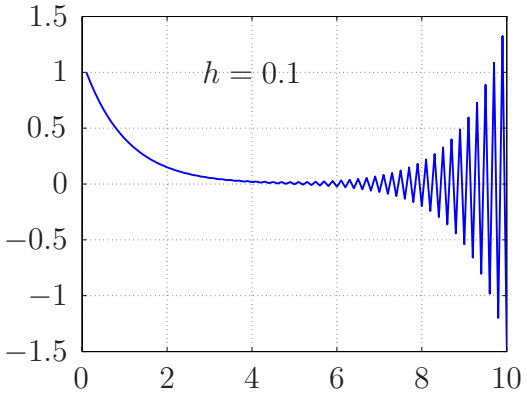
Ein Beispiel eines expliziten 2-Schrittverfahren ist die Mittelpunktsregel

$$u^h_{\ell+1} = u^h_{\ell-1} + 2hf(t^h_\ell, u^h_\ell)$$

mit den Parametern $a_1 = 1$ und $b_0 = 2$. Sie beruht auf der symmetrischen Approximation

$$u'(t) \approx \frac{u(t+h) - u(t-h)}{2h}$$

der ersten Ableitung.



Die Abbildung zeigt die numerische Lösung für das Modellproblem

$$u' = -u, \quad u(0) = 1\,,$$

wobei die exakten Startwerte

$$u^h_0 = 1, \quad u^h_1 = e^{-h}$$

verwendet wurden. Man erkennt, dass die Approximation zwar anfangs relativ genau ist, für großes t jedoch starke Oszillationen auftreten. Um dieses Phänomen zu erklären, wird die Differenzengleichung

$$u^h_{\ell+1} = u^h_{\ell-1} - 2hu^h_\ell, \quad u^h_0 = 1,\ u^h_1 = e^{-h}\,,$$

genauer untersucht. Die Lösung hat die Form

$$u^h_\ell = c_1\lambda_1^\ell + c_2\lambda_2^\ell$$

mit

$$\lambda_1 = -h + \sqrt{1+h^2} = e^{-h} + O(h^3), \quad \lambda_2 = -h - \sqrt{1+h^2} = -e^h + O(h^3)$$

den Nullstellen des charakteristischen Polynoms

$$\lambda^2 + 2h\lambda - 1 = 0\,.$$

Die Koeffizienten c_k lassen sich mit Hilfe eines Computer-Algebra-Systems aus den Anfangsbedingungen bestimmen:

$$c_1 = 1 - \frac{1}{12}h^3 + O(h^4), \quad c_2 = \frac{1}{12}h^3 + O(h^4)\,.$$

Während der erste Term in der Darstellung von u^h_ℓ der exakten Lösung entspricht,

$$c_1\lambda_1^\ell = e^{-T}(1 + O(h^2T)), \quad T = \ell h\,,$$

hat der zweite ein völlig falsches Verhalten:

$$c_2\lambda_2^\ell = (-1)^\ell \frac{h^3}{12} e^T (1 + O(h^2 T)) .$$

Aufgrund des Faktors $h^3/12$ ist dieser Term zwar anfänglich klein. Für großes T dominiert er aber die abklingende Komponente. Solche sogenannten parasitären Lösungskomponenten sind typisch für Mehrschrittverfahren. Sie können Stabilitätsprobleme verursachen, die sich nur durch sehr kleine Schrittweiten vermeiden lassen.

2.2 Ordnung eines Mehrschrittverfahrens

Ein lineares n-Schrittverfahren mit Parametern

$$u_{\ell+1}^h = \sum_{k=0}^{n-1} a_k u_{\ell-k}^h + h \sum_{k=-1}^{n-1} b_k f(t_{\ell-k}^h, u_{\ell-k}^h)$$

zur Approximation eines Differentialgleichungssystems

$$u' = f(t, u)$$

hat die Ordnung m, wenn für glattes f und glatte Lösungen u

$$\Delta(t,h) = \frac{1}{h} \sum_{k=0}^{n-1} a_k u(t - kh) + \sum_{k=-1}^{n-1} b_k f(t - kh, u(t - kh)) = O(h^m)$$

mit $a_{-1} = -1$. Die Ordnung des Diskretisierungsfehlers Δ, der beim Einsetzen einer exakten Lösung in die Differenzengleichung entsteht, kann durch Bedingungen an die Parameter a_k und b_k charakterisiert werden. Es müssen die Ordnungsbedinungen

$$\sum_{k=-1}^{n-1} a_k k^j = j \sum_{k=-1}^{n-1} b_k k^{j-1}$$

für $j = 0, \ldots, m$ erfüllt und für $j = m + 1$ verletzt sein.
Hat ein Verfahren mindestens die Ordnung 1, so wird es als konsistent bezeichnet. Anders als bei Einschrittverfahren ist Konsistenz nur notwendig aber nicht hinreichend für die Konvergenz eines Mehrschrittverfahrens.

Beweis:

Zur Herleitung der Ordnungsbedingungen ersetzt man in dem Ausdruck für den Diskretisierungsfehler $f(t - kh, u(t - kh))$ durch $u'(t - kh)$ und substituiert die Taylor-Entwicklungen

$$\begin{aligned} u(t - kh) &= \sum_{j\geq 0} \frac{u^{(j)}(t)}{j!} (-kh)^j \\ u'(t - kh) &= \sum_{j\geq 0} \frac{u^{(j+1)}(t)}{j!} (-kh)^j . \end{aligned}$$

Damit ist der Koeffizient von h^{j-1} in der Darstellung von $\Delta(t,h)$ gleich

$$\sum_k a_k \frac{u^{(j)}(t)}{j!}\,(-k)^j + \sum_k b_k \frac{u^{(j)}(t)}{(j-1)!}\,(-k)^{j-1}\,.$$

Nach Ausklammern des Faktors $(-1)^j u^{(j)}(t)/j!$ erhält man durch Nullsetzen dieses Ausdrucks die j-te Ordnungsbedingung.

Beispiel:

Als Beispiel wird ein explizites 2-Schrittverfahren betrachtet:

$$u^h_{\ell+1} = a_0 u^h_\ell + a_1 u^h_{\ell-1} + h\left(b_0 f(t^h_\ell, u^h_\ell) + b_1 f(t^h_{\ell-1}, u^h_{\ell-1})\right)\,.$$

Für Konsistenz, d.h. die minimale Ordnung 1, müssen die ersten beiden Ordnungsbedingungen gelten:

$$\begin{array}{lrcl} \underline{j=0}: & -1+a_0+a_1 & = & 0 \\ \underline{j=1}: & 1+0+a_1 & = & b_0+b_1 \end{array}$$

Eine mögliche Wahl sind die Parameter der Mittelpunktsregel:

$$a_0 = 0,\ a_1 = 1,\ b_0 = 2,\ b_1 = 0\,.$$

Diese erfüllen auch die nächste Ordnungsbedingung

$$\underline{j=2}: \quad -1+0+a_1 \;=\; 2(0+b_1)\,,$$

d.h. der Fehler der Mittelpunktsregel hat die Ordnung $O(h^2)$.
Durch die vierte Bedingung

$$\underline{j=3}: \quad 1+0+a_1 \;=\; 3(0+b_1)$$

werden die Parameter für ein explizites 2-Schrittverfahren eindeutig festgelegt:

$$a_0 = -4,\ a_1 = 5,\ b_0 = 4,\ b_1 = 2\,.$$

Das entsprechende Verfahren hat die Ordnung 3, ist jedoch nicht stabil. Es existieren exponentiell wachsende parasitäre Lösungen, so dass die Methode nicht angewendet werden sollte.

2.3 Adams-Bashforth-Verfahren

Das Adams-Bashforth-Verfahren zur Approximation des Differentialgleichungssystems

$$u' = f(t,u)$$

basiert auf der Identität

$$u(t+h) = u(t) + \int_t^{t+h} f(s,u(s))\,ds\,.$$

Man approximiert den Integrand durch ein (vektorwertiges) Polynom p vom Grad $< n$, das an den Punkten $s = t^h_\ell, \ldots, t^h_{\ell-n+1}$ $(t^h_\ell = t_0 + \ell h)$ interpoliert.

Mit der Lagrange-Darstellung von p erhält man ein n-Schrittverfahren der Ordnung n:

$$u^h_{\ell+1} = u^h_\ell + h\sum_{k=0}^{n-1} b_k f(t^h_{\ell-k}, u^h_{\ell-k})$$

mit

$$b_k = \int_0^1 \prod_{j=0, j\neq k}^{n-1} \frac{s+j}{j-k}\, ds\,.$$

Die Koeffizienten der Verfahren bis zur Ordnung $n = 6$ sind in der folgenden Tabelle angegeben.

n	b_0	b_1	b_2	b_3	b_4	b_5
1	1					
2	$\frac{3}{2}$	$-\frac{1}{2}$				
3	$\frac{23}{12}$	$-\frac{16}{12}$	$\frac{5}{12}$			
4	$\frac{55}{24}$	$-\frac{59}{24}$	$\frac{37}{24}$	$-\frac{8}{24}$		
5	$\frac{1901}{720}$	$-\frac{2774}{720}$	$\frac{2616}{720}$	$-\frac{1274}{720}$	$\frac{251}{720}$	
6	$\frac{4277}{1440}$	$-\frac{7923}{1440}$	$\frac{9982}{1440}$	$-\frac{7298}{1440}$	$\frac{2877}{1440}$	$-\frac{475}{1440}$

Beweis:

Mit

$$g_\ell = f(t^h_\ell, u^h_\ell)$$

ist die Lagrange-Form des Interpolationspolynoms

$$p(s) = \sum_{k=0}^{n-1} g_{\ell-k} q_k(s), \quad q_k(s) = \prod_{j=0, j\neq k}^{n-1} \frac{s - (t^h_\ell - jh)}{(t^h_\ell - kh) - (t^h_\ell - jh)}\,.$$

Damit erhält man die Approximation

$$u^h_{\ell+1} = u^h_\ell + \sum_k g_{\ell-k} \int_{t^h_\ell}^{t^h_{\ell+1}} q_k(s)\, ds\,.$$

Mit der Substitution

$$s = t^h_\ell + h\tilde{s}, \quad ds = hd\tilde{s}$$

folgt die Formel für b_k.
Zur Bestimmung des Diskretisierungsfehlers benutzt man die Newton-Darstellung des Interpolationsfehlers:

$$\underbrace{f(s, u(s))}_{u'(s)} - p(s) = \frac{u^{(n+1)}(\tau)}{n!}(s-t)(s-t+h)\cdots(s-t+(n-1)h)$$

mit $\tau \in [t-(n-1)h, t]$. Damit erhält man für die Norm des Diskretisierungsfehlers

$$\Delta(t,h) = \frac{u(t+h)-u(t)}{h} - \frac{1}{h}\int_t^{t+h} p(s)\, ds = \frac{1}{h}\int_t^{t+h} f(s,u(s)) - p(s)\, ds$$

die Abschätzung

$$\|\Delta(t,h)\| \leq \frac{1}{h}\,(t+h-t)\,\frac{c}{n!}\,h(2h)\cdots(nh)$$

mit c einer Schranke für $\|u^{(n+1)}(\tau)\|$.

Beispiel:

Das Adams-Bashforth-Verfahren der Ordnung 2 basiert auf linearer Interpolation der Ableitungen $g^h_{\ell-k} = f(t^h_{\ell-k}, u^h_{\ell-k})$:

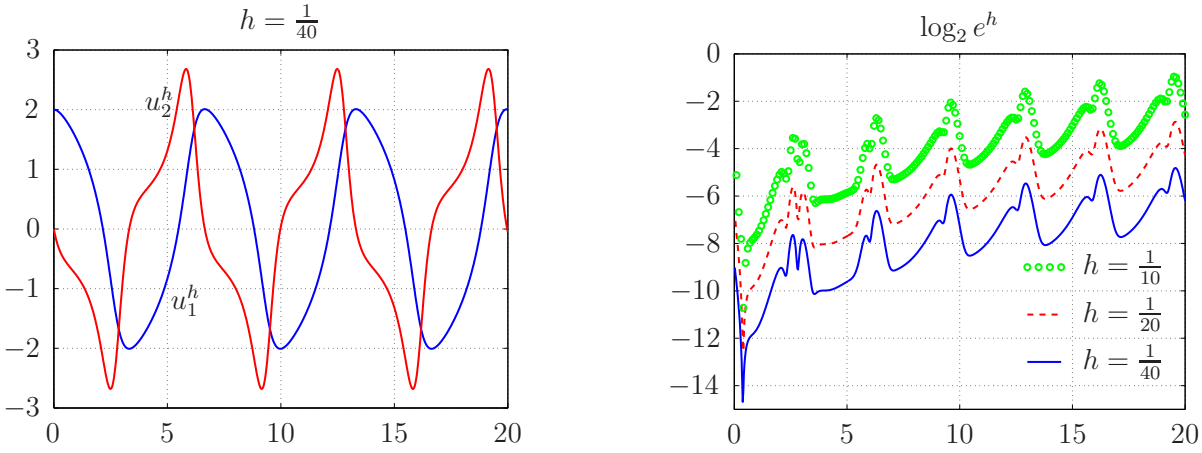
$$p(s) = g_\ell\,\frac{s-t_{\ell-1}}{h} + g_{\ell-1}\,\frac{s-t_\ell}{-h}\,.$$

Damit ist

$$\int_{t^h_\ell}^{t^h_{\ell+1}} p(s)\,ds = h\left(\frac{3}{2}g^h_\ell - \frac{1}{2}g^h_{\ell-1}\right).$$

Der zusätzlich zu dem Anfangswert $u^h_0 = u(t_0)$ benötigte Startwert $u^h_1 \approx u(t^h_1)$ kann durch Taylor-Entwicklung approximiert werden:

$$u^h_1 = u^h_0 + hf(t_0, u^h_0) = u(t^h_1) + O(h^2)\,.$$



Die Abbildung zeigt die numerische Lösung für van der Pol's Differentialgleichung

$$\begin{aligned} u_1' &= u_2 \\ u_2' &= \left(1-u_1^2\right)u_2 - u_1\,. \end{aligned}$$

Der logarithmische Fehlerplot für eine Folge sukzessive halbierter Schrittweiten bestätigt die Konvergenzordnung.

2.4 Adams-Moulton-Verfahren

Das Adams-Moulton-Verfahren zur Approximation des Differentialgleichungssystems

$$u' = f(t,u)$$

ist eine implizite Variante des Adams-Bashforth-Verfahrens. Es basiert ebenfalls auf der Identität

$$u(t+h) = u(t) + \int_t^{t+h} f(s,u(s))\,ds\,.$$

Man approximiert den Integrand durch ein (vektorwertiges) Polynom p vom Grad $< n$, das an den Punkte $s = t^h_{\ell+1}, \ldots, t^h_{\ell-n+1}$ ($t^h_\ell = t_0 + \ell h$) interpoliert. Zusätzlich zu den beim Adams-Bashforth-Verfahren verwendeten Daten wird der noch unbekannte Wert $u^h_{\ell+1}$ in die in die Interpolation mit einbezogen.
Mit der Lagrange-Darstellung von p erhält man ein n-Schrittverfahren der Ordnung $n+1$:

$$u^h_{\ell+1} = u^h_\ell + h \sum_{k=-1}^{n-1} b_k f(t^h_{\ell-k}, u^h_{\ell-k})$$

mit

$$b_k = \int_0^1 \prod_{j=-1, j\neq k}^{n-1} \frac{s+j}{j-k}\, ds\,.$$

Die Koeffizienten der Verfahren bis zur Ordnung $n = 6$ sind in der folgenden Tabelle angegeben.

n	b_{-1}	b_0	b_1	b_2	b_3	b_4	b_5
1	$\frac{1}{2}$	$\frac{1}{2}$					
2	$\frac{5}{12}$	$\frac{8}{12}$	$-\frac{1}{12}$				
3	$\frac{9}{24}$	$\frac{19}{24}$	$-\frac{5}{24}$	$\frac{1}{24}$			
4	$\frac{251}{720}$	$\frac{646}{720}$	$-\frac{264}{720}$	$\frac{106}{720}$	$-\frac{19}{720}$		
5	$\frac{475}{1440}$	$\frac{1427}{1440}$	$-\frac{798}{1440}$	$\frac{482}{1440}$	$-\frac{173}{1440}$	$\frac{27}{1440}$	
6	$\frac{19087}{60480}$	$\frac{65112}{60480}$	$-\frac{46461}{60480}$	$\frac{37504}{60480}$	$-\frac{20211}{60480}$	$\frac{6312}{60480}$	$-\frac{863}{60480}$

Da beim Adams-Moulton-Verfahren in jedem Schritt ein nichtlineares Gleichungssystem zur Bestimmung von $u^h_{\ell+1}$ gelöst werden muss, wird das Verfahren üblicherweise in Verbindung mit einem expliziten Mehrschrittverfahren zur Schätzung des gesuchten Wertes verwendet. Der Mehraufwand wird im allgemeinen durch bessere Stabilitätseigenschaften kompensiert.

Beweis:

Mit

$$g_\ell = f(t^h_\ell, u^h_\ell)$$

ist die Lagrange-Form des Interpolationspolynoms

$$p(s) = \sum_{k=-1}^{n-1} g_{\ell-k} q_k(s), \quad q_k(s) = \prod_{j=-1, j\neq k}^{n-1} \frac{s - (t^h_\ell - jh)}{(t^h_\ell - kh) - (t^h_\ell - jh)}\,.$$

Damit erhält man die Approximation

$$u^h_{\ell+1} = u^h_\ell + \sum_k g_{\ell-k} \int_{t^h_\ell}^{t^h_{\ell+1}} q_k(s)\, ds\,.$$

Mit der Substitution

$$s = t^h_\ell + h\tilde{s}, \quad ds = h d\tilde{s}$$

folgt die Formel für b_k.

Zur Bestimmung des Diskretisierungsfehlers benutzt man die Newton-Darstellung des Interpolationsfehlers:

$$\underbrace{f(s,u(s))}_{u'(s)} - p(s) = \frac{u^{(n+2)}(\tau)}{(n+1)!}(s-(t-h))(s-t)\cdots(s-t+(n-1)h))$$

mit $\tau \in [t-(n-1)h, t+h]$. Damit erhält man für die Norm des Diskretisierungsfehlers

$$\Delta(t,h) = \frac{u(t+h)-u(t)}{h} - \frac{1}{h}\int_t^{t+h} p(s)\,ds = \frac{1}{h}\int_t^{t+h} f(s,u(s)) - p(s)\,ds$$

die Abschätzung

$$\|\Delta(t,h)\| \le \frac{1}{h}\,(t+h-t)\,\frac{c}{(n+1)!}\,h\cdot h\cdot(2h)\cdots(nh)$$

mit c einer Schranke für $\|u^{(n+2)}(\tau)\|$.

2.5 BDF-Verfahren

Das BDF-Verfahren (backward-differentiation-formula) zur Lösung eines Differentialgleichungssystems

$$u' = f(t,u)$$

ist ein implizites lineares Mehrschrittverfahren. Dabei interpoliert man die Approximationen $u^h_{\ell-k} \approx u(t^h_{\ell-k})$, $t = -1,\ldots,n-1$, durch ein Polynom p vom Grad $\le n$ und fordert, dass p das Differentialgleichungssystem im Punkt $t^h_{\ell+1}$ erfüllt:

$$p'(t^h_{\ell+1}) = f(t^h_{\ell+1}, u^h_{\ell+1})\,.$$

Damit hat ein Schritt des Verfahrens die Form

$$u^h_{\ell+1} = \sum_{k=0}^{n-1} a_k u^h_{\ell-k} + hb_{-1} f(t^h_{\ell+1}, u^h_{\ell+1})$$

mit

$$b_{-1} = 1/p'_{-1}(1), \quad a_k = -p'_k(1)b_{-1}$$

und p_k den Lagrange-Polynomen

$$p_k(s) = \prod_{j=-1, j\ne k}^{n-1} \frac{s+j}{j-k}\,.$$

Das n-Schritt-BDF-Verfahren hat die Ordnung n. Allerdings sind die Verfahren nur bis zur Ordnung 6 stabil. Für $n > 6$ existieren exponentiell wachsende parasitäre Lösungen, so dass diese Verfahren nicht zur numerischen Approximation verwendet werden können. Die Parameter der stabilen BDF-Verfahren können der folgenden Tabelle entnommen werden.

n	b_{-1}	a_0	a_1	a_2	a_3	a_4	a_5
2	$\frac{2}{3}$	$\frac{4}{3}$	$-\frac{1}{3}$				
3	$\frac{6}{11}$	$\frac{18}{11}$	$-\frac{9}{11}$	$\frac{2}{11}$			
4	$\frac{12}{25}$	$\frac{48}{25}$	$-\frac{36}{25}$	$\frac{16}{25}$	$-\frac{3}{25}$		
5	$\frac{60}{137}$	$\frac{300}{137}$	$-\frac{300}{137}$	$\frac{200}{137}$	$-\frac{75}{137}$	$\frac{12}{137}$	
6	$\frac{60}{147}$	$\frac{360}{147}$	$-\frac{450}{147}$	$\frac{400}{147}$	$-\frac{225}{147}$	$\frac{72}{147}$	$-\frac{10}{147}$

Beweis:

Die Lagrange-Form des Interpolationspolynoms ist

$$p(t) = \sum_{k=-1}^{n-1} u^h_{\ell-k} p_k((t - t^h_\ell)/h)$$

mit

$$p_k(s) = \prod_{j=-1, j\neq k}^{n-1} \frac{s+j}{j-k}\,.$$

Damit ist die Gleichung

$$p'(t^h_{\ell+1}) = f(t^h_{\ell+1}, u^h_{\ell+1})$$

äquivalent zu

$$u^h_{\ell+1}\,(p'_{-1}(1)/h) = -\sum_{k=0}^{n-1} u^h_{\ell-k}(p'_k(1)/h)\; + hf(t^h_{\ell+1}, u^h_{\ell+1})\,,$$

woraus sich die Koeffizienten $a_0, \ldots, a_{n-1}$ and b_{-1} ablesen lassen.
Bezeichnet q das Interpolationspolynom der exakten Lösungswerte, so hat der Diskretisierungsfehler die Form

$$\Delta(t,h) = \frac{1}{p'_{-1}(1)}\,\left[q'(t+h) - u'(t+h)\right]\,.$$

Mit der Newton-Darstellung des Interpolationsfehlers ist

$$[\ldots] = \frac{d}{ds}\left(\Delta(s, t+h, t, \ldots, t-(n-1)h)u\, w(s)\right)_{|s=t+h}$$

mit $w(s) = (s-(t+h))\cdots(s-(t-(n-1)h))$. Anwendung der Produktregel unter Berücksichtigung von

$$\frac{d}{ds}\Delta(s, \tau_1, \ldots, \tau_m)u = \Delta(s, s, \tau_1, \ldots, \tau_m)u, \quad \Delta(\sigma_0, \ldots, \sigma_m)u = \frac{1}{m!}u^{(m)}(\xi)$$

sowie

$$w(t+h) = O(h^{m+1}), \quad w'(t+h) = O(h^m)$$

zeigt, dass Δ die behauptete Ordnung n besitzt.

Beispiel:

Als Beispiel wird das BDF-Verfahren der Ordnung 2 betrachtet. Benutzt man die Newton-Form des Interpolationspolynoms, so ist

$$p(t) = u^h_{\ell-1} + \frac{u^h_\ell - u^h_{\ell-1}}{h}\,(t - t^h_{\ell-1}) + \frac{u^h_{\ell+1} - 2u^h_\ell + u^h_{\ell-1}}{2h^2}\,(t - t^h_{\ell-1})(t - t^h_\ell)$$

und

$$p'(t_{\ell+1}) = \frac{u^h_\ell - u^h_{\ell-1}}{h} + \frac{u^h_{\ell+1} - 2u^h_\ell + u^h_{\ell-1}}{2h^2}\,(2h+h)\,.$$

Somit folgt

$$p'(t^h_{\ell+1} = f(t^h_{\ell+1}, u^h_{\ell+1}) \Leftrightarrow \frac{3}{2}u^h_{\ell+1} - 2u^h_\ell + 12u^h_{\ell-1} = hf(t^h_{\ell+1}, u^h_{\ell+1})\,.$$

Durch Auflösen nach $u^h_{\ell+1}$ ergibt sich die Standardform des Mehrschrittverfahrens mit den Koeffizienten

$$a_0 = 4/3,\; a_1 = -1/3, \quad b_{-1} = 2/3\,.$$

2.6 Prädiktor-Korrektor-Verfahren

Das Prädiktor-Korrektor-Verfahren ist eine mögliche Realisierung eines Zeitschritts

$$u(t_\ell^h) \approx u_\ell^h \to u_{\ell+1}^h \approx u(t_{\ell+1}^h)$$

eines impliziten linearen Mehrschrittverfahrens zur Approximation eines Differentialgleichungssystems

$$u' = f(t,u)\,.$$

Für ein n-Schritt-Verfahren der Ordnung m wird dabei zunächst ein Schritt eines expliziten n-Schritt-Verfahrens der Ordnung mindestens $m-1$ durchgeführt:

$$v_{\ell+1}^h = \sum_{k=0}^{n-1} a_k^p u_{\ell-k}^h + h\sum_{k=0}^{n-1} b_k^p f(t_{\ell-k}^h, u_{\ell-k}^h)\,.$$

Dann wird die Approximation des impliziten Verfahrens durch einen Schritt einer Fixpunktiteration approximiert:

$$u_{\ell+1}^h = \sum_{k=0}^{n-1} a_k^c u_{\ell-k}^h + h\sum_{k=0}^{n-1} b_k^c f(t_{\ell-k}^h, u_{\ell-k}^h) + hb_{-1}^c f(t_{\ell+1}^h, v_{\ell+1}^h)\,.$$

Die Ordnung des impliziten Verfahrens bleibt bei dieser Korrektur erhalten.

Beweis:

Der Diskretisierungsfehler eines Prädiktor-Korrektor-Schrittes ist

$$\Delta(t,h) = \frac{1}{h}\sum_{k=0}^{n-1} a_k^c u(t-kh) - \sum_{k=0}^{n-1} b_k^c u'(t-kh) - b_{-1}^c f(t-h, v(t+h))$$

mit $a_{-1} = -1$ und

$$v(t+h) = \sum_{k=0}^{n-1} a_k^p u(t-kh) + h\sum_{k=0}^{n-1} b_k^p u'(t-kh)\,.$$

Da der Prädiktor mindestens die Ordnung $m-1$ hat, ist

$$v(t+h) = u(t+h) + \delta, \quad \|\delta\| = O(h^m)\,.$$

Substituiert man dies in den Ausdruck für $\Delta(t,h)$ und benutzt, dass

$$f(t+h, u(t+h)+\delta) = f(t+h, u(t+h)) + O(\|\delta\|)\,.$$

so folgt, dass $\Delta(t,h)$ sich von dem Diskretisierungsfehler des impliziten Verfahrens mit Parametern a_k^c und b_k^c nur um $O(h^m)$ unterscheidet, also die gleiche Ordnung hat.

Beispiel:

Ein typisches Prädiktor-Korrektor-Verfahren erhält man durch Kombination der n-Schritt-Adams-Bashforth- und Adams-Moulton-Verfahren. Beispielsweise hat für $n = 2$ ein Zeitschritt die Form

(P) $v^h_{\ell+1} = u^h_\ell + \frac{h}{2}(3f(t^h_\ell, u^h_\ell) - f(t^h_{\ell-1}, u^h_{\ell-1}))$

(C) $u^h_{\ell+1} = u^h_\ell + \frac{h}{12}(5f(t^h_{\ell+1}, v^h_{\ell+1}) + 8f(t^h_\ell, u^h_\ell) - f(t^h_{\ell-1}, u^h_{\ell-1}))$.

Um die Ordnung zu testen, wird das Modellproblem

$$u' = u$$

betrachtet. Mit $u^h_\ell = 1$, $u^h_{\ell-1} = e^{-h}$ erhält man

$$v^h_{\ell+1} = 1 + \frac{h}{2}(3 \cdot 1 - e^{-h}) = 1 + h + \frac{h^2}{2} + O(h^3)$$

und

$$u^h_{\ell+1} = 1 + \frac{h}{12}(5(1 + h + h^2/2 + O(h^3)) + 8 - e^{-h}) = 1 + h + \frac{h^2}{2} + \frac{h^3}{6} + O(h^4)\,.$$

Der lokale Fehler zur Lösung $u(t) = \exp(t - t^h_\ell)$ hat also die Ordnung 4 in Übereinstimmung mit der Ordnung 3 des Diskretisierungsfehlers des 2-Schritt-Adams-Moulton-Verfahrens.

Kapitel 3

Stabilität und Schrittweitensteuerung

3.1 Matrix-Norm und Spektralradius

Für den Spektralradius $\varrho(A)$ einer quadratischen Matrix A gilt

$$\varrho(A) \leq \|A\|$$

für jede einer Vektornorm zugeordnete Matrix-Norm.
Die umgekehrte Ungleichung gilt im allgemeinen nicht. Es existiert jedoch für alle $\varepsilon > 0$ eine Vektornorm $\| \ \|_\varepsilon$, so dass

$$\|A\|_\varepsilon \leq \varrho(A) + \varepsilon \, .$$

Stimmt für alle Eigenwerte λ von A mit $|\lambda| = \varrho(A)$ die geometrische mit der algebraischen Vielfachheit überein, so ist $\varepsilon = 0$ möglich, d.h. $\|A\|_\varepsilon = \varrho(A)$.

Beweis:

Zunächst transformiert man A auf Jordan-Normalform:

$$J = Q^{-1} A Q \, .$$

Durch eine weitere Ähnlichkeitstransformation mit der Diagonalmatrix $S = \operatorname{diag}(1, \varepsilon, \varepsilon^2, \ldots)$, kann man die Einsen auf der Nebendiagonalen auf ε skalieren. Setzt man nun

$$\|x\|_\varepsilon = \|S^{-1} Q^{-1} x\|_\infty \, ,$$

so folgt

$$\|A\|_\varepsilon = \max_{y \neq 0} \frac{\|S^{-1} Q^{-1} A y\|_\infty}{S^{-1} Q^{-1} y\|_\infty} = \max_{x \neq 0} \frac{\|Bx\|_\infty}{\|x\|_\infty} \, .$$

Die konstruierte Matrix-Norm ist also die Zeilensummennorm von B:

$$\|A\|_\varepsilon = \|B\|_\infty = \max_\lambda |\lambda + \sigma_\lambda \varepsilon| \leq \varrho(A) + \varepsilon$$

mit $\sigma_\lambda \in \{0, 1\}$. Dabei ist σ_λ nur dann gleich 1, wenn in der Jordan-Form neben dem Eigenwert λ eine Eins steht. Insbesondere ist also σ_λ gleich 0, wenn die geometrische mit der algebraischen Vielfachheit von λ übereinstimmt.
Es kann ε so klein gewählt werden, dass für $|\lambda| < \varrho(A)$ auch $|\lambda + \varepsilon| \leq \varrho(A)$ ist. Besitzt A keine Eigenwerte λ mit $|\lambda| = \varrho(A)$ und $\sigma_\lambda = 1$ stimmt dann das Maximum mit $\varrho(A)$ überein.

3.2 Stabilität linearer Mehrschrittverfahren

Ein lineares n-Schrittverfahren

$$u^h_{\ell+1} = \sum_{k=0}^{n-1} a_k u^h_{\ell-k} + h \sum_{k=-1}^{n-1} b_k f(t^h_{\ell-k}, u^h_{\ell-k})$$

zur Approximation eines Differentialgleichungssystems

$$u' = f(t,u)$$

bezeichnet man als stabil, wenn numerische Lösungen eines homgenen linearen Differentialgleichungssystems

$$u' = Q(t)u$$

gleichmäßig bzgl. der Schrittweite h beschränkt sind. Genauer gilt

$$\|u^h_\ell\| \le c \max_{0\le k<n} \|u^h_k\|, \quad \ell h \le T\,,$$

wobei die Konstante c nur von T und $\max_{t_0\le t\le T} \|Q(t)\|$ abhängt. Aus dieser Abschätzung folgt insbesondere, dass die numerische Lösung stetig von den Startwerten $u^h_0, \ldots, u^h_{n-1}$ abhängt.
Notwending und hinreichend für Stabilität ist die sogenannte Wurzelbedingung. Nullstellen λ des charakteristischen Polynoms

$$p(\lambda) = \lambda^n - \sum_{k=0}^{n-1} a_k \lambda^{n-1-k}$$

müssen in der komplexen Einheitskreisscheibe liegen und, wenn sie Betrag Eins haben, einfach sein:

- $|\lambda| \le 1$;
- $|\lambda| = 1 \implies p'(\lambda) \neq 0$.

Beweis:

Es sind beide Richtungen der Äquivalenz

$$\text{Stabilität} \Leftrightarrow \text{Wurzelbedingung}$$

zu zeigen.
$\Longrightarrow$: Den Nullstellen des charakteristischen Polynoms entsprechen Lösungen der homogenen Differenzengleichung

$$u^h_{\ell+1} = a_0 u^h_\ell + \cdots + a_{n-1} u^h_{\ell-n+1}\,.$$

Für eine einfache Nullstelle λ ist

$$u^h_\ell = \lambda^\ell$$

eine Lösung und für eine mehrfache Nullstelle zusätzlich

$$u^h_\ell = \ell\lambda^\ell\,.$$

Beide Lösungen sind numerische Approximationen der trivialen Differentialgleichung

$$u' = 0$$

mit den Startwerten

$$\ell^\nu \lambda^\ell, \quad \ell = 0, \dots, n-1\,,$$

und $\nu = 0$ oder 1. Ist die Wurzelbedingung verletzt, so existiert eine Nullstelle λ mit $|\lambda| > 1$ für $\nu = 0$ oder mit $|\lambda| = 1$ für $\nu = 1$. Da die entsprechende Lösung der Differenzengleichung nicht von h abhängt, wird in beiden Fällen das Verhältnis

$$|u_\ell^h| / \max_{0 \le k < n} |u_k^h|$$

für $\ell h = T = 1$ und $h \to 0$ beliebig groß. Die Stabilitätsabschätzung kann also nicht gelten.
$\Leftarrow$: Um die Stabilitätsabschätzung zu zeigen, muss man das Wachstum der Lösung der Differenzengleichung

$$u_{\ell+1}^h = \sum_{k=0}^{n-1} a_k u_{\ell-k}^h + h \sum_{k=-1}^{n-1} b_k A(t_{\ell-k}^h) u_{\ell-k}^h$$

analysieren. Dazu fasst man n aufeinanderfolgende Approximationen zu einer $d \times n$-Matrix zusammen:

$$V_\ell = (u_\ell^h, \dots, u_{\ell-n+1}^h)\,.$$

Mit

$$M = \begin{pmatrix} a_0 & 1 & & \\ a_1 & & \ddots & \\ \vdots & & & 1 \\ a_{n-1} & 0 & \cdots & 0 \end{pmatrix}$$

hat dann die Rekursion die Form

$$V_{\ell+1} = V_\ell M + hG(V_{\ell+1}, V_\ell)\,,$$

wobei G für $\ell h \le T$ der Abschätzung

$$\|G(V_{\ell+1}, V_\ell)\| \le c_G(\|V_{\ell+1}\| + \|V_\ell\|)$$

genügt. Schranken für V_ℓ hängen entscheidend von der Größe der Matrix M ab. Da $(-1)^n p(\lambda)$ das charakteristische Polynom von M ist, erfüllen die Eigenwerte die Wurzelbedingung. Mit Hilfe der Jordan-Normalform von M lässt sich damit eine Norm $\|\ \|_*$ konstruieren, für die in der induzierten Matrix-Norm

$$\|M\|_* \le 1$$

gilt. Für $c_G h \le 1/2$ folgt nun

$$\|V_{\ell+1}\| \le \frac{1 + c_G h}{1 - c_G h} \|V_\ell\|_* \le (1 + 3c_G h)\|V_\ell\|_*\,.$$

Durch Iteration dieser Abschätzung erhält man

$$\|V_\ell\|_* \le (1 + 3c_G h)^\ell \|V_0\|_* \le \exp(3c_G \ell h)\,.$$

Die numerische Lösung lässt sich also durch die Startwerte beschränken:

$$\|V_\ell\|_* \le c\|V_0\|$$

mit $c = \exp(3c_G T)$.

Beispiel:

Für das Adams-Bashforth-Verfahren der Ordnung n ist

$$a_0 = 1,\ a_1 = \cdots = a_{n-1} = 0\,.$$

Folglich ist das charakteristische Polynom

$$p(\lambda) = \lambda^n - \lambda^{n-1}$$

und hat die Nullstellen $\lambda = 1$ und $\lambda = 0$ mit Vielfachheit $n-1$. Das n-Schritt-Adams-Moulton-Verfahren hat dasselbe charakteristische Polynom, denn die Koeffizienten a_k sind unverändert, nur $a_0 = 1$ ist ungleich Null. Beide Verfahren sind also für jede Ordnung stabil.

Beispiel:

Das BDF-Verfahren der Ordnung 2,

$$u^h_{\ell+1} = -\frac{1}{3}u^h_{\ell-1} + \frac{4}{3}u^h_\ell - \frac{2}{3}hf\left(t^h_{\ell+1}, u^h_{\ell+1}\right)$$

hat das charakteristische Polynom

$$p\left(\lambda\right) = \lambda^2 - \frac{4}{3}\lambda + \frac{1}{3}$$

mit den Nullstellen $\lambda = 1$ und $\lambda = \frac{1}{3}$, ist also stabil.
Die Nullstellen für die BDF-Verfahren höherer Ordnung können der folgenden Tabelle entnommen werden. Man erkennt, dass die Wurzelbedingung bis zur Ordnung 6 erfüllt ist.

Ordnung	Nullstellen
1	1
2	$\frac{1}{3}$, 1
3	1, 0.3182 ± 0.2839i
4	0.3815, 1, 0.2693 ± 0.4920i
5	1, 0.3848 ± 0.1621i, 0.2100 ± 0.6769i
6	0.4061, 1, 0.3762 ± 0.2885i, 0.1453 ± 0.8511i

3.3 Konvergenz linearer Mehrschrittverfahren

Erfüllt ein n-Schrittverfahren der Ordnung $m \geq 1$ das Wurzelkriterium, und approximieren die Startwerte $u^h_\ell \approx u(t^h_\ell)$ die Lösung des Anfangswertproblems

$$u' = f(t,u), \quad u(t_0) = u_0\,,$$

mit der Ordnung m,

$$\|u^h_\ell - u(t^h_\ell)\| = O(h^m), \quad \ell = 0, \ldots, n-1\,,$$

so hat der globale Fehler ebenfalls die Ordnung m:

$$\|u^h_\ell - u(t^h_\ell)\| = O(h^m), \quad 0 \leq \ell h \leq T\,.$$

Dabei wird vorausgesetzt, dass f glatt ist, die Lösung u auf dem Intervall $[t_0, t_0 + T]$ existiert und h genügend klein ist.

Beweis:

Zunächst wird das n-Schritt-Verfahren

$$u^h_{\ell+1} = \sum_{k=0}^{n-1} a_k u^h_{\ell-k} + h \sum_{k=-1}^{n-1} b_k f(t^h_{\ell-k}, u^h_{\ell-k})$$

als Einschrittverfahren umgeschrieben. Dazu fasst man n aufeinanderfolgende Approximationen zu einer $d \times n$-Matrix zusammen:

$$V_\ell = (u^h_\ell, \ldots, u^h_{\ell-n+1}) .$$

Mit

$$M = \begin{pmatrix} a_0 & 1 & & \\ a_1 & & \ddots & \\ \vdots & & & 1 \\ a_{n-1} & 0 & \cdots & 0 \end{pmatrix}$$

erhält man

$$V_{\ell+1} = V_\ell M + h \left[\sum_{k=-1}^{n-1} b_k f(t^h_{\ell-k}, u^h_{\ell-k}),\, 0,\, \ldots,\, 0 \right] .$$

Die genaue Form der $d \times n$-Matrix $h[\ldots]$ wird nicht benötigt. Sie kann deshalb in der Form $hG(t^h_\ell, h, V^h_{\ell+1}, V^h_\ell)$ abgekürzt werden. Setzt man

$$W^h_\ell = \big(u(t^h_\ell), \ldots, u(t^h_{\ell-n+1})\big) ,$$

so gilt nach Definition des Diskretisierungsfehlers

$$W^h_{\ell+1} = W^h_\ell M + hG(t^h_\ell, h, W^h_{\ell+1}, W^h_\ell) + h\left(\Delta(t^h_\ell, h), 0, \ldots, 0\right) .$$

Subtrahiert man die entsprechende Gleichung für $V^h_{\ell+1}$, so folgt aufgrund der Glattheit von G für den Fehler $E^h_\ell = W^h_\ell - V^h_\ell$

$$\|E^h_{\ell+1}\| \le \|E^h_\ell\|\|M\| + hc_G(\|E^h_{\ell+1}\| + \|E^h_\ell\|) + hc_\Delta h^m .$$

Nach Konstruktion sind die Eigenwerte von M die Nullstellen des charakteristischen Polynoms

$$p(\lambda) = \lambda^n - a_0\lambda^{n-1} - \cdots - a_{n-1}$$

des n-Schritt-Verfahrens. Aufgrund der Wurzelbedingung existiert folglich eine Norm $\|\ \|_*$, so dass in der zugeordneten Matrix-Norm $\|M\|_* \le 1$ gilt. Für $hc_G \le 1/2$ ist somit

$$\|E^h_{\ell+1}\|_* \le \frac{1 + c_G h}{1 - c_G h}\|E^h_\ell\|_* + 2hc_\Delta h^m .$$

Der Bruch lässt sich durch $1 + 3c_G h$ abschätzen. Mit $c = \max(3c_G, 2c_\Delta)$ folgt dann durch Iteration der Ungleichung

$$\begin{aligned} \|E^h_\ell\|_* &= (1+ch)\|E^h_{\ell-1}\|_* + ch^{m+1} \\ &= (1+ch)^2\|E^h_{\ell-2}\|_* + (1+ch)ch^{m+1} + ch^{m+1} \\ &= \cdots \\ &= (1+ch)^{\ell-n+1}\|E^h_{n-1}\|_* + \frac{(1+ch)^{\ell-n+2} - 1}{ch} ch^{m+1} . \end{aligned}$$

Mit Hilfe der Ungleichung $(1 + ch) \le \exp(ch)$ erhält man schließlich

$$\|E^h_\ell\|_* \le \exp(c\ell h)(\|E^h_{n-1}\|_* + ch^m) .$$

Da $\ell h \le T$ und sich $\|E^h_{n-1}\|_*$ durch den maximalen Fehler der Startwerte abschätzen lässt, hat der Ausdruck die gewünschte Ordnung $O(h^m)$.

3.4 Stabilitätsgebiete von Einschritt- und Mehrschrittverfahren

Eine numerische Lösung $u_\ell^h \approx u(t_\ell^h)$, $t_\ell^h = t_0 + \ell h$, eines Differentialgleichungssystems

$$u' = f(t,u)$$

sollte das gleiche qualitative Verhalten wie die exakte Lösung haben. Besonders kritisch ist dabei die Approximation schnell abklingender Lösungskomponenten, wie z.B. für das Modellproblem

$$u' = \lambda u, \quad \operatorname{Re}\lambda < 0\,.$$

Dies motiviert die folgende Definition. Man bezeichnet die Menge aller komplexen Zahlen $\omega = h\lambda$, für die die numerische Lösung des Modellproblems für $\ell \to \infty$ gegen Null streben, als Stabilitätsgebiet Ω des Verfahrens.
Die Relevanz des Modellproblems beruht auf der Linearisierung

$$f \approx f_t\,(t - t_0) + f_u\,(u(t) - u(t_0))\,.$$

Aufgrund der Lösungsstruktur für lineare Differentialgleichungssysteme spielen die Eigenwerte der Jacobi-Matrix f_u die Rolle des Parameters λ in dem Modellproblem. Als Voraussetzung für ein richtiges qualitatives Verhalten der Lösung sollte die Schrittweite h mindestens so klein gewählt werden, dass $\lambda h \in \Omega$ für alle Eigenwerte mit negativem Realteil.

Beispiel:

Zur Bestimmung des Stabilitätsgebietes Ω für ein n-stufiges Runge-Kutta-Verfahren mit Parametermatrix

$$\left(\begin{array}{c|c} A & c \\ \hline b^t & \end{array}\right)$$

wird ein Schritt

$$u(t) \approx v \to w \approx u(t+h)$$

betrachtet. Für das Modellproblem berechnen sich die Hilfsgrößen z_k aus dem linearen Gleichungssystem

$$z_k = \lambda(v + h(a_{k,1}z_1 + \cdots + a_{k,n}z_n)), \quad k = 1,\ldots,n\,,$$

und

$$w = v + h(b_1 z_1 + \cdots + b_n z_n)\,.$$

Nach der Cramerschen Regel ist

$$z_k = \lambda v \frac{p_k(\omega)}{q(\omega)}, \quad \omega = h\lambda\,,$$

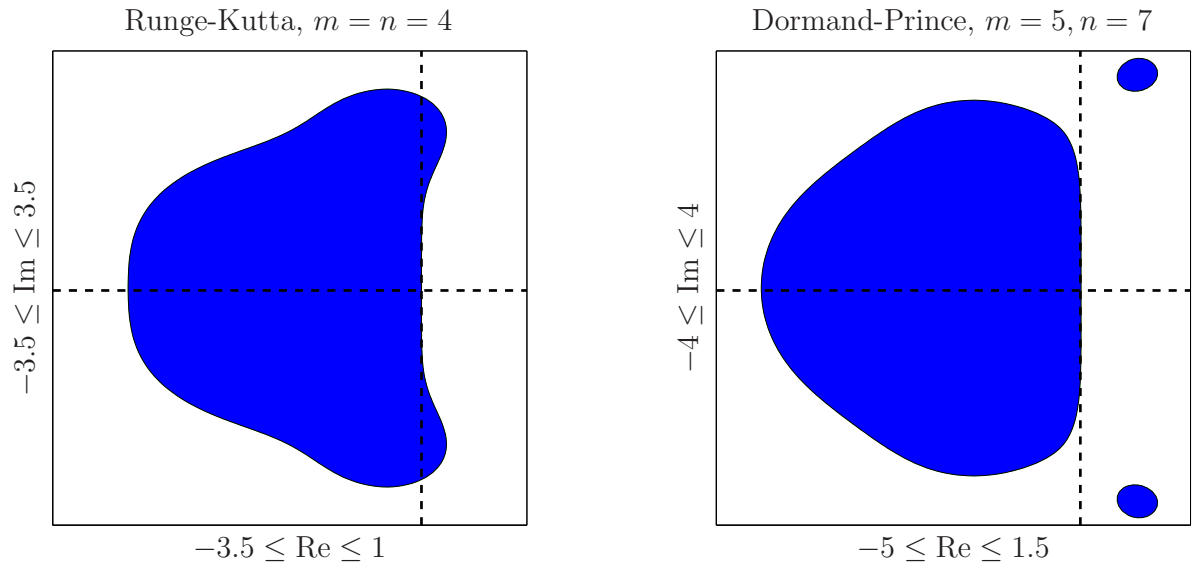
mit Polynomen p_k und q vom Grad $\le n-1$ bzw $\le n$. Folglich ist

$$w = vr(\omega)$$

mit r einer rationalen Funktion mit Zähler- und Nennergrad $\le n$. Damit gilt

$$\Omega : |r(\omega)| < 1\,.$$

Bei expliziten Runge-Kutta-Verfahren können die Hilfsgrößen durch sukzessives Einsetzen bestimmt werden, ohne dass ein lineares Gleichungssystem gelöst werden muss. In diesem Fall ist $q = 1$ und die Funktion r ein Polynom.



Die Abbildung zeigt die Stabilitätsgebiete der beiden gebräuchlichsten expliziten Runge-Kutta-Verfahren. Für alle Gauß-Runge-Kutta-Verfahren gilt

$$\Omega : \mathrm{Re}\,\omega < 0\,.$$

Es besteht damit für diese impliziten Verfahren optimaler Ordnung keine stabilitätsbedingte Einschränkung an die Schrittweite. Für alle $h > 0$ hat die numerische Lösung das gleiche qualitative Verhalten wie die exakte Lösung des Modellproblems.

Beispiel:

Für das Modellproblem

$$u' = \lambda u$$

hat ein lineares n-Schritt-Verfahren die Form

$$\sum_{k=-1}^{n-1} (a_k + \omega b_k) u^h_{\ell-k} = 0, \quad \omega = h\lambda\,,$$

mit $a_{-1} = -1$. Numerische Lösungen erfüllen genau dann $|u^h_\ell| \to 0$, $\ell \to \infty$, für beliebige Startwerte $u^h_0, \ldots, u^h_{n-1}$, wenn alle Nullstellen des Polynoms

$$p(z) + \omega q(z) = \sum_{k=-1}^{n-1} (a_k + \omega b_k) z^{n-1-k}$$

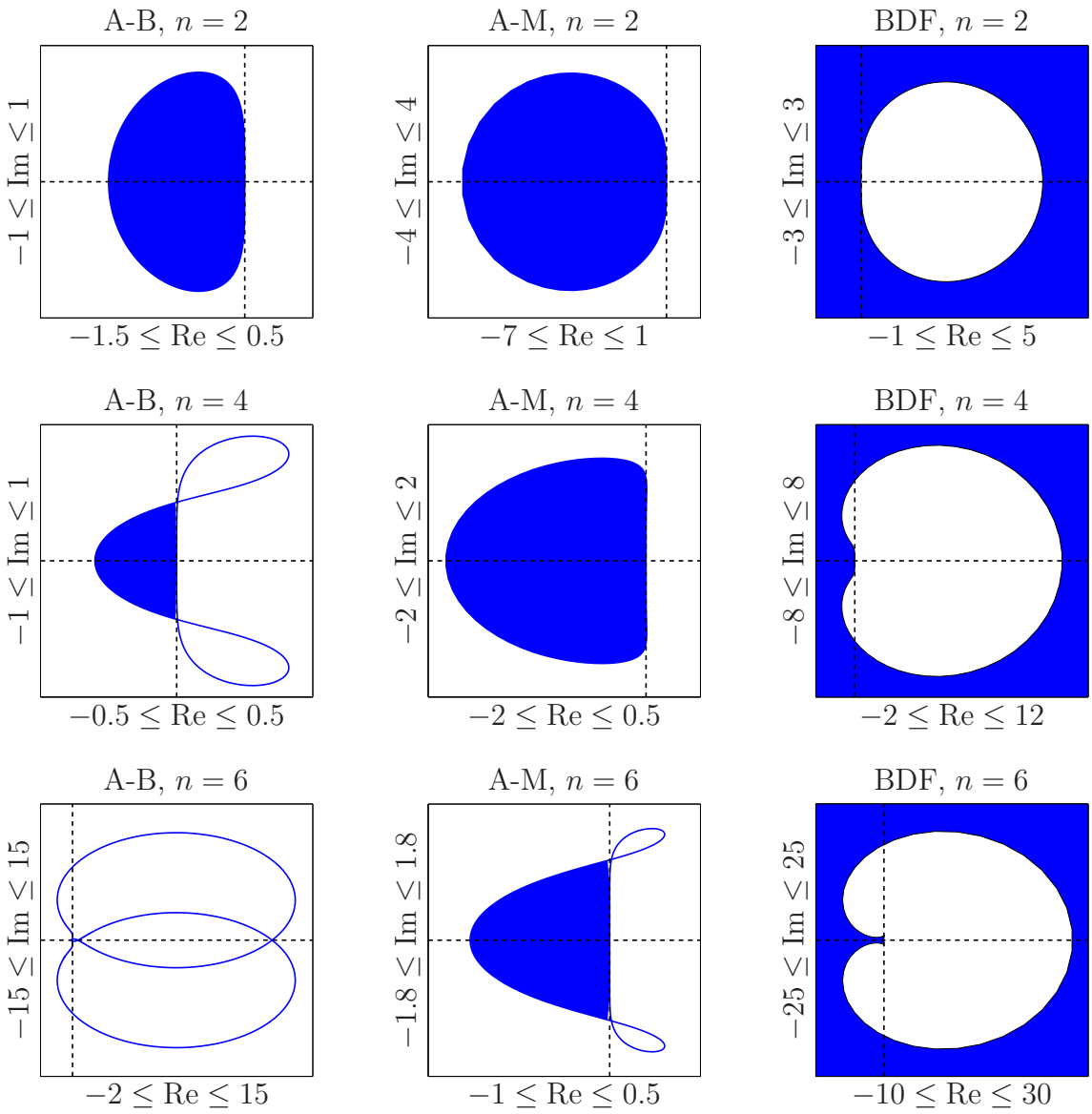
Betrag < 1 haben:

$$\Omega : p(z) = -\omega q(z) \implies |z| < 1\,.$$

Der Rand von Ω wird somit von Teilen der Kurve

$$\frac{p(e^{\mathrm{i}t})}{q(e^{\mathrm{i}t})}, \quad 0 \leq t < 2\pi\,,$$

gebildet.



Die Abbildung zeigt Stabilitätsgebiete für einige der Adams-Bashforth- und Adams-Moulton-Verfahren. Es fällt auf, dass Ω für die impliziten Verfahren größer ist. Die aufwändigere Implementierung wird damit durch geringere stabilitätsbedingte Einschränkungen an die Schrittweite kompensiert.

3.5 Schrittweitensteuerung

Für eine effiziente Berechnung einer numerischen Lösung einer Differentialgleichung sollte die verwendete Schrittweite h an das Problem angepasst werden. Einerseits muss die Schrittweite klein genug sein, um eine vorgegebene Toleranz einzuhalten, andererseits sollte die Schrittzahl möglichst klein bleiben.
Damit der globale Fehler die Größenordnung $\varepsilon(t_1 - t_0)$ hat, sollte für den lokalen Fehler

$$\|d_{\mathrm{loc}}\| \leq \varepsilon h$$

gelten. Zu einem Verfahren der Ordnung m ist daher die optimale Schrittweite

$$h_{\mathrm{opt}} = h(\varepsilon h / \|d_{\mathrm{loc}}\|)^{1/m} .$$

Wird der lokale Fehler d_{loc} geschätzt, sollte der Ungenauigkeit der Schätzung durch einen Faktor $h_{\text{factor}} < 1$ Rechnung getragen werden. Darüber hinaus empfiehlt es sich, Vergrößerungen der Schrittweite nur langsam durchzuführen.

Beweis:

Für ein Verfahren der Ordnung m hat der lokale Fehler bei einer Schrittweite h die Ordnung $O(h^{m+1})$ und sollte bei einer optimalen Schrittweite h_{opt} gleich $\varepsilon h_{\text{opt}}$ sein. Damit ist

$$\|d_{\text{loc}}\| = ch^{m+1}\,, \quad \varepsilon h_{\text{opt}} = ch_{\text{opt}}^{m+1}\,.$$

Löst man die erste Gleichung nach c auf,

$$c = \|d_{\text{loc}}\|/h^{m+1}\,,$$

setzt den Ausdruck in die zweite Gleichung ein,

$$\varepsilon h_{\text{opt}} = \|d_{\text{loc}}\|h_{\text{opt}}^{m+1}/h^{m+1}\,,$$

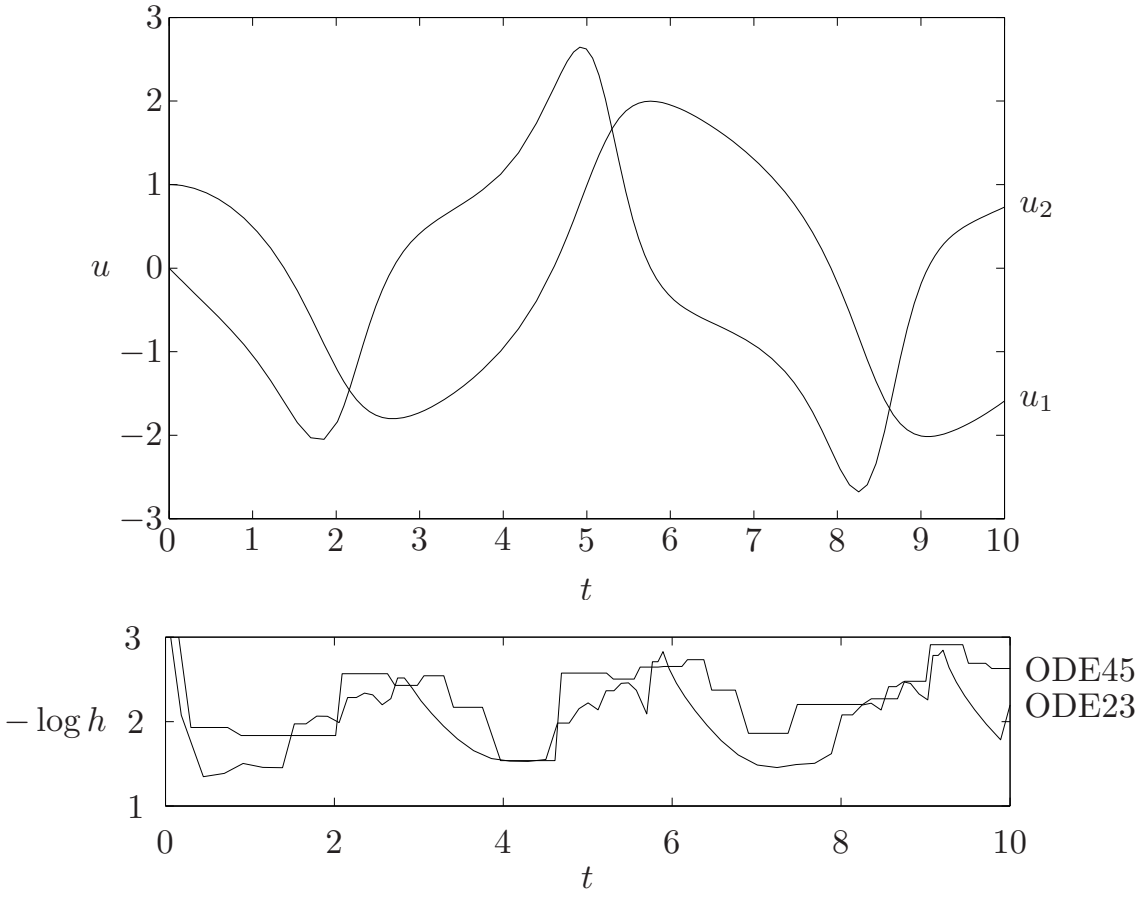
so erhält man durch Auflösen nach h_{opt} die angegebene Form.

Beispiel:

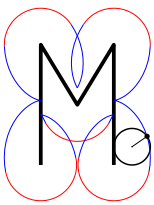
Zur Illustration der Schrittweitensteuerung wird die Van der Polsche Differentialgleichung

$$\begin{aligned} u_1' &= u_2 \\ u_2' &= (1-u_1^2)u_2 - u_1 \end{aligned}$$

mit den Matlab-Programmen ODE23 und ODE45 gelöst.



Das erste Bild zeigt die Lösung der $(u_1(t), u_2(t))^{\mathrm{t}}$. Im zweiten Bild sind die Schrittweiten in Abhängigkeit von der Zeit logarithmisch dargestellt. Erwartungsgemäß wird die Schrittweite in Bereichen mit großen Ableitungen der Lösung reduziert. Durch diese lokale Anpassung wird eine höhere Effizienz als mit konstanter Schrittweite erreicht.



http://www.mathematik-online.org/