

Berechnung des charakteristischen Exponenten der endlichen Hillschen Differentialgleichung durch Numerische Integration

E. Wagenführer und H. Lang

Universität Regensburg, Fachbereich Mathematik,
Postfach 397, D-8400 Regensburg 2, Bundesrepublik Deutschland

On Calculating the Characteristic Exponent of the Finite Hill Differential Equation by Numerical Integration

Summary. The characteristic exponent ν of the finite Hill equation

$$(*) \quad y''(x) + \left(\lambda + 2 \sum_{k=1}^l t_k \cos(2kx) \right) y(x) = 0$$

satisfies the equations

$$\cos(\pi\nu) = 2y_1\left(\frac{\pi}{2}\right)y_2'\left(\frac{\pi}{2}\right) - 1 = 2y_2\left(\frac{\pi}{2}\right)y_1'\left(\frac{\pi}{2}\right) + 1,$$

where y_1, y_2 are the canonical fundamental solutions of (*). For calculating y_1, y_2 the Taylor expansion method of a high order p ($10 \leq p \leq 40$) turns out to be the best of all known methods of numerical integration. In this paper the Taylor method for solving (*) is formulated, an extensive error analysis – including the rounding errors – is performed. If the parameters in (*) are not too large, the computed error bounds will be rather realistic.

Subject Classifications. AMS(MOS); 65L05; CR: 5.16.

Einleitung

Der charakteristische Exponent der Mathieuschen Differentialgleichung läßt sich nach Schäfke, Ebert, Groh und Schmidt [6–8] über geeignete unendliche Determinanten und damit über dreigliedrige Rekursionen berechnen. Diese Verfahren sind wesentlich günstiger als das in [7], S. 7 zum Vergleich betrachtete Runge-Kutta-Verfahren. Von Mennicken [4] und Wagenführer [10] wurde die Determinantenmethode verallgemeinert bzw. in der Konvergenz verbessert.

Die vorliegende Arbeit behandelt für den allgemeinen Fall der endlichen Hillschen Differentialgleichung die Berechnung des charakteristischen Exponenten mittels Numerischer Integration. Die Arbeit basiert auf umfangreichen Untersuchungen, in denen die wichtigsten heute bekannten Integrationsverfahren praktisch durchgerechnet wurden: hierbei haben sich Taylor-Verfahren hoher Ordnung als besonders geeignet erwiesen.

In Abschnitt 1 wird das Taylor-Verfahren für die endliche Hillsche Differentialgleichung – unter Berücksichtigung programmtechnischer Aspekte – formuliert. Die Abschnitte 2–4 befassen sich mit einer vollständigen Fehleranalyse. In Abschnitt 2 wird der lokale Diskretisierungsfehler abgeschätzt; hieraus wird eine Methode abgeleitet, mit der die Ordnung des Verfahrens geeignet festgelegt werden kann. Abschnitt 3 enthält Algorithmen, mit denen der lokale Rundungsfehler abzuschätzen ist: wegen der relativ großen Schrittweiten ist die Rundungsfehleranalyse nicht trivial. Abschnitt 4 behandelt die globalen Fehler, wobei die sonst übliche Fehleranalyse wesentlich verfeinert wird. Abschnitt 5 schließlich befaßt sich mit der praktischen Erfahrung: er bringt Beispiele, an denen unter anderem die Qualität der angegebenen Fehleranalyse demonstriert wird, und Vergleiche mit anderen Verfahren.

Die Autoren danken Herrn Prof. Dr. R. Mennicken für die Anregung zu den vorliegenden Untersuchungen.

1. Beschreibung des Verfahrens

Der charakteristische Exponent ν der endlichen Hillschen Differentialgleichung

$$(1.1) \quad y''(x) + \left(\lambda + 2 \sum_{k=1}^l t_k \cos(2kx) \right) y(x) = 0 \quad (\lambda, t_k \in \mathbb{C})$$

ist durch die Existenz einer Lösung $y \neq 0$ von (1.1) mit der Eigenschaft

$$y(x + \pi) = e^{i\pi\nu} y(x) \quad (x \in \mathbb{R})$$

definiert. Bekanntlich ist mit ν auch $-\nu$ charakteristischer Exponent; sind ferner y_1, y_2 die Lösungen von (1.1) mit den Anfangsbedingungen

$$\begin{pmatrix} y_1(0) \\ y_1'(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} =: e_1, \quad \begin{pmatrix} y_2(0) \\ y_2'(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} =: e_2,$$

so gelten die Beziehungen

$$(1.2) \quad \begin{aligned} \cos(\pi\nu) &= 2y_1\left(\frac{\pi}{2}\right)y_2'\left(\frac{\pi}{2}\right) - 1 \\ &= 2y_2\left(\frac{\pi}{2}\right)y_1'\left(\frac{\pi}{2}\right) + 1 \end{aligned}$$

– vgl. [3], S. 100–103 –. Welche der beiden Gln. (1.2) sich zur Berechnung von ν besser eignet, hängt vom Wert der rechten Seiten ab. Liegt dieser beispielsweise

nahe bei 1, so sollte man die zweite Gleichung in der stabilisierten Form

$$(1.2') \quad \sin^2\left(\frac{\pi}{2} v\right) = -y_2\left(\frac{\pi}{2}\right) y_1'\left(\frac{\pi}{2}\right)$$

verwenden. – Somit ist die Berechnung von v zurückgeführt auf die Lösung von Anfangswertaufgaben der Gestalt

$$(1.3) \quad y''(x) = g(x) y(x), \quad y(0) = y_0, \quad y'(0) = y_0'$$

mit $y_0, y_0' \in \mathbb{R}$ sowie

$$(1.4) \quad g(x) = -\left(\lambda + \sum_{k=1}^l (2t_k) \cos(2kx)\right).$$

Zur Lösung y von (1.3) definieren wir

$$(1.5) \quad z(x) := \begin{pmatrix} y(x) \\ y'(x) \end{pmatrix} \quad (x \in \mathbb{R}).$$

Dann ist z ebenso wie y in ganz \mathbb{R} beliebig oft differenzierbar; nach dem Satz von Taylor gilt für $x \in \mathbb{R}, h > 0, p \in \mathbb{N}$

$$(1.6) \quad z(x+h) = z(x) + \sum_{\mu=1}^p \frac{h^\mu}{\mu!} z^{(\mu)}(x) + R_p(x, h)$$

mit dem Restglied

$$(1.7) \quad R_p(x, h) = \frac{h^{p+1}}{p!} \int_0^1 (1-t)^p z^{(p+1)}(x+th) dt.$$

Auf Grund von (1.3) erfüllen die Komponenten der $z^{(\mu)}(x)$ die Gleichungen

$$(1.8) \quad y^{(\mu)}(x) = \sum_{\kappa=0}^{\mu-2} \binom{\mu-2}{\kappa} g^{(\mu-2-\kappa)}(x) y^{(\kappa)}(x) \quad (\mu = 2, 3, 4, \dots).$$

Dies begründet das nun zu beschreibende Taylor-Verfahren. Wir geben uns eine Schrittweite

$$h = \frac{\pi}{2N} \quad (N \in \mathbb{N}, > 0)$$

vor und wählen demgemäß die Stützstellen

$$x_n = nh \quad (n = 0, 1, 2, \dots, N).$$

Wir ermitteln, beispielsweise so wie in Abschnitt 2 beschrieben, eine geeignete Ordnung $p (\geq 2)$ des Verfahrens. Dann berechnen wir als Näherungen für die $z(x_n)$

$$v_n = \begin{pmatrix} u_n \\ u_n' \end{pmatrix}$$

vermöge der Rekursion

$$(1.9) \quad \begin{aligned} v_0 = z_0 &:= \begin{pmatrix} y_0 \\ y'_0 \end{pmatrix}, \\ v_{n+1} &= v_n + \sum_{\mu=1}^p \frac{h^\mu}{\mu!} v_n^{(\mu)} \quad (n=0, 1, \dots, N-1), \end{aligned}$$

in der die

$$v_n^{(\mu)} := \begin{pmatrix} u_n^{(\mu)} \\ u_n^{(\mu+1)} \end{pmatrix}$$

wiederum rekursiv durch

$$(1.10) \quad \begin{aligned} u_n^{(0)} &= u_n, & u_n^{(1)} &= u'_n, \\ u_n^{(\mu)} &= \sum_{\kappa=0}^{\mu-2} \binom{\mu-2}{\kappa} g^{(\mu-2-\kappa)}(x_n) u_n^{(\kappa)} \quad (\mu=2, 3, \dots) \end{aligned}$$

definiert sind. Bekanntlich ist für $\kappa \geq 1$

$$(1.11) \quad g^{(\kappa)}(x) = \begin{cases} (-1)^{\frac{\kappa-1}{2}} \sum_{k=1}^l 2t_k (2k)^\kappa \sin(2kx) & (\kappa \text{ ungerade}), \\ (-1)^{\frac{\kappa-2}{2}} \sum_{k=1}^l 2t_k (2k)^\kappa \cos(2kx) & (\kappa \text{ gerade, } \geq 2); \end{cases}$$

demnach müssen die Werte $\sin x, \cos x$ für

$$x = 2kx_n = \frac{2kn}{N} \cdot \frac{\pi}{2} \quad (k=0, 1, \dots, l; n=0, 1, \dots, N)$$

bestimmt werden. Auf Grund der elementaren Eigenschaften der trigonometrischen Funktionen lassen sich zu jedem $(k, n) \in \mathbb{N} \times \mathbb{N}$ mittels ganzzahliger Operationen $k', k'' \in \{0, 1, \dots, N\}$ berechnen, so daß

$$(1.12) \quad \sin\left(\frac{2kn}{N} \cdot \frac{\pi}{2}\right) = \pm \sin\left(k' \frac{\pi}{2N}\right), \quad \cos\left(\frac{2kn}{N} \cdot \frac{\pi}{2}\right) = \pm \sin\left(k'' \frac{\pi}{2N}\right)$$

gilt. Insgesamt werden daher nur $N-1$ Sinus-Werte gebraucht.

Da mit wachsendem κ die Koeffizienten der $g^{(\kappa)}(x)$ sehr groß werden, empfiehlt es sich, direkt die Ausdrücke

$$\frac{h^\kappa}{\kappa!} g^{(\kappa)}(x_n)$$

zu berechnen. Für die in (1.9) auftretenden

$$\frac{h^\mu}{\mu!} u_n^{(\mu)} \quad (\mu=0, 1, 2, \dots), \quad \frac{h^{\mu-1}}{(\mu-1)!} u_n^{(\mu)} \quad (\mu=1, 2, \dots)$$

verwendet man gemäß (1.10) die Rekursionen

$$(1.13) \quad \begin{aligned} u_n^{(0)} &= u_n, & u_n^{(1)} &= u'_n, & h u_n^{(1)} &= h \cdot u'_n; \\ \frac{h^{\mu-1}}{(\mu-1)!} u_n^{(\mu)} &= \frac{h}{\mu-1} \left[\sum_{\kappa=0}^{\mu-2} \frac{h^{\mu-2-\kappa}}{(\mu-2-\kappa)!} g^{(\mu-2-\kappa)}(x_n) \left(\frac{h^\kappa}{\kappa!} u_n^{(\kappa)} \right) \right], \\ \frac{h^\mu}{\mu!} u_n^{(\mu)} &= \frac{h}{\mu} \left(\frac{h^{\mu-1}}{(\mu-1)!} u_n^{(\mu)} \right) \quad (\mu=2, 3, \dots). \end{aligned}$$

2. Abschätzung des lokalen Diskretisierungsfehlers

Zunächst suchen wir eine a priori Abschätzung der Lösung von (1.3). – Wir bezeichnen im Fall $\operatorname{Re}(\lambda) > 0$

$$\begin{aligned} \rho &:= \sqrt{\operatorname{Re}(\lambda)}, \\ h(x) &:= \sum_{k=1}^l (2t_k) \cos(2kx) + i \operatorname{Im}(\lambda), \\ L &:= |\operatorname{Im}(\lambda)| + \sum_{k=1}^l |2t_k|, \\ V(x) &:= \begin{pmatrix} \cos(\rho x) & \frac{1}{\rho} \sin(\rho x) \\ -\rho \sin(\rho x) & \cos(\rho x) \end{pmatrix}, \end{aligned}$$

ferner im Fall $\operatorname{Re}(\lambda) \leq 0$

$$\begin{aligned} \rho &:= 0, \\ h(x) &:= -g(x), \\ L &:= |\lambda| + \sum_{k=1}^l |2t_k|, \\ V(x) &:= \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Die letztere Definition von $V(x)$ stimmt mit der obigen überein, wenn man für $\rho=0$

$$\frac{1}{\rho} \sin(\rho x) =: x$$

setzt; dies werden wir der Einfachheit halber im Folgenden tun.

Gemäß (1.3) löst z das Differentialgleichungssystem

$$(2.1) \quad \begin{aligned} z'(x) + \begin{pmatrix} 0 & -1 \\ \rho^2 & 0 \end{pmatrix} z(x) &= \begin{pmatrix} 0 & 0 \\ -h(x) & 0 \end{pmatrix} z(x), \\ z(0) &= z_0, \end{aligned}$$

ferner ist $V(x)$ Fundamentalmatrix des zugehörigen „homogenen“ Systems. Auf Grund der Lösungsformel für inhomogene Differentialgleichungssysteme erfüllt daher z die Integralgleichung

$$(2.2) \quad z(x) = \int_0^x \begin{pmatrix} -\frac{1}{\rho} \sin(\rho(x-\xi)) h(\xi) & 0 \\ -\cos(\rho(x-\xi)) h(\xi) & 0 \end{pmatrix} z(\xi) d\xi + V(x) z_0$$

– vgl. auch [2], S. 271 –. Wir betrachten (2.2) als Fixpunktproblem

$$z = T(z)$$

im Raum $C_0\left(\left[0, \frac{\pi}{2}\right], \mathbb{C}^2\right)$, dem wir die Norm

$$(2.3) \quad \|w\| := \max \left\{ |w_1(x)|, |w_2(x)| : x \in \left[0, \frac{\pi}{2}\right] \right\} \quad \left(w(x) = \begin{pmatrix} w_1(x) \\ w_2(x) \end{pmatrix} \right)$$

beziehungsweise im Fall $\rho > 0$ wahlweise die Norm

$$(2.3') \quad \|w\|_\rho := \max \left\{ |w_1(x)|, \frac{1}{\rho} |w_2(x)| : x \in \left[0, \frac{\pi}{2}\right] \right\}$$

zugrunde legen. Bezüglich der hiervon erzeugten Metriken, die wir d bzw. d_ρ nennen, ist $C_0\left(\left[0, \frac{\pi}{2}\right], \mathbb{C}^2\right)$ jeweils ein vollständiger metrischer Raum. Wir notieren im Sinne der in [5], Bd. 1, S. 76 gegebenen Definition den

(2.4) **Hilfssatz.** *T ist beschränkter Operator; bezüglich der Metrik d gilt*

$$(2.5) \quad \sum_{\kappa=0}^{\infty} \|T^\kappa\| \leq 1 + \sqrt{L} \sinh\left(\sqrt{L} \frac{\pi}{2}\right),$$

ferner hat man im Fall $\rho > 0$ bezüglich d_ρ

$$(2.5') \quad \sum_{\kappa=0}^{\infty} \|T^\kappa\|_\rho \leq \exp\left(\frac{L}{\rho} \cdot \frac{\pi}{2}\right).$$

Beweis. Es seien $w, \tilde{w} \in C_0\left(\left[0, \frac{\pi}{2}\right], \mathbb{C}^2\right)$ vorgegeben. Wir bezeichnen für $\kappa \in \mathbb{N}$

$$w^\kappa = \begin{pmatrix} w_1^\kappa \\ w_2^\kappa \end{pmatrix} := T^\kappa(w), \quad \tilde{w}^\kappa = \begin{pmatrix} \tilde{w}_1^\kappa \\ \tilde{w}_2^\kappa \end{pmatrix} := T^\kappa(\tilde{w})$$

und behaupten für $\kappa \geq 1$, $x \in \left[0, \frac{\pi}{2}\right]$

$$(2.6) \quad \begin{aligned} |w_1^\kappa(x) - \tilde{w}_1^\kappa(x)| &\leq \frac{L^\kappa}{(2\kappa)!} x^{2\kappa} \|w - \tilde{w}\|, \\ |w_2^\kappa(x) - \tilde{w}_2^\kappa(x)| &\leq \frac{L^\kappa}{(2\kappa-1)!} x^{2\kappa-1} \|w - \tilde{w}\| \end{aligned}$$

sowie im Fall $\rho > 0$

$$(2.6) \quad \begin{aligned} |w_1^\kappa(x) - \tilde{w}_1^\kappa(x)| &\leq \frac{L^\kappa}{\kappa!} \frac{1}{\rho^\kappa} x^\kappa \|w - \tilde{w}\|_\rho, \\ \frac{1}{\rho} |w_2^\kappa(x) - \tilde{w}_2^\kappa(x)| &\leq \frac{L^\kappa}{\kappa!} \frac{1}{\rho^\kappa} x^\kappa \|w - \tilde{w}\|_\rho. \end{aligned}$$

Den Nachweis von (2.6) bzw. (2.6') erbringt man durch Induktion über κ . Dabei benutzt man für $\kappa \geq 0$ die Beziehungen

$$w^{\kappa+1}(x) - \tilde{w}^{\kappa+1}(x) = \int_0^x \begin{pmatrix} -\frac{1}{\rho} \sin(\rho(x-\xi)) h(\xi) (w_1^\kappa(\xi) - \tilde{w}_1^\kappa(\xi)) \\ -\cos(\rho(x-\xi)) h(\xi) (w_1^\kappa(\xi) - \tilde{w}_1^\kappa(\xi)) \end{pmatrix} d\xi$$

und beachtet

$$|h(\xi)| \leq L, \quad |\cos(\rho(x-\xi))| \leq 1;$$

ferner zieht man zum Beweis von (2.6) die Eigenschaft

$$\left| \frac{1}{\rho} \sin(\rho(x-\xi)) \right| \leq x - \xi$$

bzw. im Fall (2.6') die Ungleichung

$$\left| \frac{1}{\rho} \sin(\rho(x-\xi)) \right| \leq \frac{1}{\rho}$$

heran. – Wegen

$$\frac{\pi}{2} \frac{1}{2^\kappa} \leq 1$$

folgt aus (2.6) bezüglich der Metrik d

$$\|T^\kappa\| \leq \frac{L^\kappa}{(2\kappa-1)!} \left(\frac{\pi}{2}\right)^{2\kappa-1} \quad (\kappa \geq 1),$$

ferner schließt man aus (2.6') bezüglich d_ρ unmittelbar

$$\|T^\kappa\|_\rho \leq \frac{1}{\kappa!} \left(\frac{L}{\rho} \cdot \frac{\pi}{2}\right)^\kappa,$$

womit die Eigenschaften (2.5), (2.5') aufgezeigt sind.

Der Operator T erfüllt die Voraussetzungen des Banachschen Fixpunktsatzes, wie er beispielsweise in [5], Bd. 2, S. 82 notiert ist. Demnach gilt für den

Fixpunkt z von T bei beliebigem $\tilde{z} \in C_0\left(\left[0, \frac{\pi}{2}\right], \mathbb{C}^2\right)$ die Abschätzung

$$(2.7) \quad d_{(\rho)}(z, \tilde{z}) \leq \left(\sum_{\kappa=0}^{\infty} \|T^\kappa\|_{(\rho)}\right) d_{(\rho)}(\tilde{z}, T(\tilde{z})).$$

Wir wenden (2.7) auf $\tilde{z}=0$ an, dann wird $T(\tilde{z})=Vz_0$; ferner interessieren nur die Werte $z_0=e_1$ oder e_2 . Offenbar gilt

$$\|Ve_1\| \leq \max\{1, |\rho|\}, \quad \|Ve_2\| \leq \frac{\pi}{2}$$

und im Fall $\rho > 0$

$$\|Ve_1\|_\rho \leq 1, \quad \|Ve_2\|_\rho \leq \frac{1}{\rho}.$$

Von den möglichen Abschätzungen, die sich hieraus zusammen mit Hilfssatz (2.4) ergeben, greifen wir nur die wichtigsten heraus:

(2.8) **Folgerung.** Für die Lösung von (1.3) mit der Anfangsbedingung $z_0=e_1$ oder e_2 gilt im Fall $\rho \leq 1$

$$(2.9) \quad \|z\| \leq \left(1 + \sqrt{L} \sinh\left(\sqrt{L} \frac{\pi}{2}\right)\right) \cdot \frac{\pi}{2} =: K$$

bzw. im Fall $\rho > 1$

$$(2.9') \quad \|z\|_\rho \leq \exp\left(\frac{L}{\rho} \frac{\pi}{2}\right) =: K_\rho.$$

Für Fehlerbetrachtungen benötigen wir mehrfach den folgenden Hilfssatz, den wir fast unmittelbar aus (1.8) bzw. (1.10) erschließen:

(2.10) **Hilfssatz.** Es sei für $x \in \mathbb{R}$

$$(2.11) \quad \begin{aligned} \alpha_0^1(x) &= 1, & \alpha_1^1(x) &= 0; & \alpha_0^2(x) &= 0, & \alpha_1^2(x) &= 1, \\ \alpha_\mu^i(x) &= \sum_{\kappa=0}^{\mu-2} \binom{\mu-2}{\kappa} g^{(\mu-2-\kappa)}(x) \alpha_\kappa^i(x) \quad (i=1, 2; \mu=2, 3, 4, \dots) \end{aligned}$$

definiert. Dann gilt für alle $\mu \in \mathbb{N}$

$$(2.12) \quad y^{(\mu)}(x) = \alpha_\mu^1(x) y(x) + \alpha_\mu^2(x) y'(x) \quad (x \in \mathbb{R}),$$

$$(2.13) \quad u_n^{(\mu)} = \alpha_\mu^1(x_n) u_n + \alpha_\mu^2(x_n) u'_n \quad (n=0, 1, \dots, N-1).$$

Um den Hilfssatz anzuwenden, definieren wir

$$(2.14) \quad F_\mu := \begin{cases} |\lambda| + \sum_{k=1}^i |2t_k| & (\mu=0), \\ \sum_{k=1}^i (2k)^\mu |2t_k| & (\mu=1, 2, 3, \dots) \end{cases}$$

sowie

$$(2.15) \quad \begin{aligned} a_0^1 &= 1, & a_1^1 &= 0; & a_0^2 &= 0, & a_1^2 &= 1, \\ a_\mu^i &= \sum_{\kappa=0}^{\mu-2} \binom{\mu-2}{\kappa} F_{\mu-2-\kappa} a_\kappa^i \quad (i=1, 2; \mu=2, 3, 4, \dots). \end{aligned}$$

Dann hat man wegen (1.11) für $\mu \in \mathbb{N}$, $x \in \mathbb{R}$ offenbar

$$|g^{(\mu)}(x)| \leq F_\mu$$

und daher weiter

$$(2.16) \quad |\alpha_\mu^i(x)| \leq a_\mu^i \quad (i=1, 2).$$

Es ergeben sich auf Grund von (2.12) für $\mu \in \mathbb{N}$, $x \in \left[0, \frac{\pi}{2}\right]$ die Abschätzungen

$$(2.17) \quad |y^{(\mu)}(x)| \leq a_\mu^1 |y(x)| + a_\mu^2 |y'(x)| \\ \leq \begin{cases} (a_\mu^1 + a_\mu^2) \|z\|, \\ (a_\mu^1 + \rho a_\mu^2) \|z\|_\rho, & \text{falls } \rho > 0. \end{cases}$$

Als *lokalen Diskretisierungsfehler* an der Stelle x_n bezeichnen wir gemäß (1.7) den Vektor

$$(2.18) \quad r_n = \begin{pmatrix} \rho_n \\ \rho'_n \end{pmatrix} := R_p(x_n, h),$$

dieser läßt sich – im Fall $z_0 = e_1$ oder e_2 – nach (2.8), (2.17) komponentenweise durch

$$(2.19) \quad \begin{pmatrix} |\rho_n| \\ |\rho'_n| \end{pmatrix} \leq r := \begin{cases} K \frac{h^{p+1}}{(p+1)!} (a_{p+1}^1 + a_{p+1}^2) & (\rho \leq 1), \\ K_\rho \frac{h^{p+1}}{(p+1)!} (a_{p+1}^1 + \rho a_{p+1}^2) & (\rho > 1) \end{cases}$$

abschätzen. – Ergänzend notieren wir den

(2.20) **Hilfssatz.** Für alle $h > 0$, $i \in \{1, 2\}$ konvergiert die Reihe

$$\sum_{\mu=0}^{\infty} \frac{h^\mu}{\mu!} a_\mu^i.$$

Beweis. Wir definieren für $h \in \mathbb{C}$

$$F(h) := |\lambda| + \sum_{k=1}^l |2t_k| e^{2kh} = \sum_{\mu=0}^{\infty} \frac{h^\mu}{\mu!} F_\mu.$$

Dann ist F in ganz \mathbb{C} holomorph; folglich sind für $i=1, 2$ die Lösungen B_i der Anfangswertaufgaben

$$(2.21) \quad B_i''(h) = F(h) B_i(h), \\ B_i(0) = a_0^i, \quad B_i'(0) = a_1^i$$

ebenfalls in ganz \mathbb{C} holomorph. Die Koeffizienten b_μ^i der Potenzreihen

$$B_i(h) = \sum_{\mu=0}^{\infty} \frac{h^\mu}{\mu!} b_\mu^i$$

genügen auf Grund von (2.21) der Rekursion (2.15), folglich hat man

$$a_\mu^i = b_\mu^i \quad (i=1, 2; \mu \in \mathbb{N}).$$

Insbesondere konvergieren für jedes $h > 0$ die rechten Seiten von (2.19) für $p \rightarrow \infty$ gegen Null. Dies kann dazu benutzt werden, um bei vorgegebener Schrittweite h die Ordnung p des Taylor-Verfahrens geeignet festzulegen. Hierzu berechnet man für $\mu=0, 1, 2, \dots$ die

$$\frac{h^\mu}{\mu!} F_\mu, \quad \frac{h^{\mu-1}}{(\mu-1)!} a_\mu^i, \quad \frac{h^\mu}{\mu!} a_\mu^i \quad (i=1, 2),$$

wobei ähnliche Rekursionen wie (1.13) auftreten, und wählt p als die kleinste Zahl, für die die rechte Seite von (2.19) in beiden Komponenten kleiner als ein vorgegebenes $\varepsilon > 0$ wird.

3. Zum lokalen Rundungsfehler

Der Einfachheit halber nehmen wir an, die Parameter $\lambda, 2t_k$ ($k=1, \dots, l$) seien reelle, dezimale Gleitkommazahlen mit t -stelliger Mantisse; das in Abschnitt 1 beschriebene Taylor-Verfahren werde mit t -stelliger Gleitkomma-Arithmetik durchgeführt. t sei genügend groß, so daß

$$\tau := \frac{1}{2} 10^{-t+1}$$

die Ungleichung

$$(3.1) \quad (4(p+1)+l)\tau \leq 0,09$$

erfüllt. Wir bezeichnen

$$\tau' := \frac{1}{1 - (4(p+1)+l)\tau} \cdot \tau \quad (\leq 1, 1\tau).$$

In der benutzten Arithmetik gelte für je zwei t -stellige Gleitkommazahlen x, y und für jede Rechenoperation $\Delta \in \{+, -, \cdot, : \}$ eine Beziehung

$$(3.2) \quad \mathbf{gl}(x \Delta y) = (x \Delta y)(1 + \eta), \quad |\eta| \leq \tau$$

und entsprechend

$$(3.3) \quad \mathbf{gl}(\sin x) = \sin x \cdot (1 + \tilde{\eta}), \quad |\tilde{\eta}| \leq \tau.$$

Unter diesen Annahmen bezeichnen wir für $n=0, 1, \dots, N$ mit

$$\tilde{v}_n = \begin{pmatrix} \tilde{u}_n \\ \tilde{u}'_n \end{pmatrix}$$

die mit Gleitkommarechnung gemäß dem Verfahren (1.9), (1.10) ermittelten numerischen Werte für v_n . Im Fall $n \leq N-1$ definieren wir

$$\tilde{u}_n^{(\mu)} \quad (\mu=0, 1, 2, \dots)$$

durch die (rundungsfehlerfrei auszuführende) Rekursion (1.10), ausgehend von den Anfangswerten $\tilde{u}_n, \tilde{u}'_n$. Entsprechend seien

$$(3.4) \quad \frac{h^\mu}{\mu!} \tilde{u}_n^{(\mu)} = \frac{h^\mu}{\mu!} (\tilde{u}_n^{(\mu)} + \varepsilon_n^{(\mu)}), \quad \frac{h^{\mu-1}}{(\mu-1)!} \hat{u}_n^{(\mu)} = \frac{h^{\mu-1}}{(\mu-1)!} (\tilde{u}_n^{(\mu)} + \varepsilon_n^{(\mu)})$$

die aus (1.13) mittels Gleitkommarechnung gewonnenen Größen, d.h.

$$(3.5) \quad \begin{aligned} \tilde{u}_n^{(0)} &= \tilde{u}_n, & h\tilde{u}_n^{(1)} &= \mathbf{gl}(h \cdot \tilde{u}'_n), & \hat{u}_n^{(1)} &= \tilde{u}'_n; \\ \frac{h^{\mu-1}}{(\mu-1)!} \hat{u}_n^{(\mu)} &= \mathbf{gl} \left\{ \frac{h}{\mu-1} \sum_{\kappa=0}^{\mu-2} \mathbf{gl} \left(\frac{h^{\mu-2-\kappa}}{(\mu-2-\kappa)!} g^{(\mu-2-\kappa)}(x_n) \right) \frac{h^\kappa}{\kappa!} \tilde{u}_n^{(\kappa)} \right\}, \\ \frac{h^\mu}{\mu!} \tilde{u}_n^{(\mu)} &= \mathbf{gl} \left(\frac{h}{\mu} \cdot \frac{h^{\mu-1}}{(\mu-1)!} \hat{u}_n^{(\mu)} \right) \quad (\mu=2, 3, \dots, p+1). \end{aligned}$$

Hierbei kennzeichnen die Ausdrücke \mathbf{gl} die zusammengesetzten Gleitkomma-Operationen in der unten näher erläuterten Reihenfolge.

Setzt man schließlich für $\mu=0, 1, 2, \dots$

$$\tilde{v}_n^{(\mu)} := \begin{pmatrix} \tilde{u}_n^{(\mu)} \\ \tilde{u}_n^{(\mu+1)} \end{pmatrix}, \quad \tilde{v}'_n^{(\mu)} := \begin{pmatrix} \tilde{u}'_n^{(\mu)} \\ \hat{u}_n^{(\mu+1)} \end{pmatrix},$$

so hat man nach Definition der \tilde{v}_n

$$(3.6) \quad \begin{aligned} \tilde{v}_0 &= z_0, \\ \tilde{v}_{n+1} &= \mathbf{gl} \left(\sum_{\mu=0}^p \frac{h^\mu}{\mu!} \tilde{v}_n^{(\mu)} \right) \quad (n=0, 1, \dots, N-1). \end{aligned}$$

Den *lokalen Rundungsfehler*

$$s_n = \begin{pmatrix} \sigma_n \\ \sigma'_n \end{pmatrix} \quad (n \in \{0, 1, \dots, N-1\})$$

definieren wir durch die Gleichung

$$(3.7) \quad \tilde{v}_{n+1} = \sum_{\mu=0}^p \frac{h^\mu}{\mu!} \tilde{v}_n^{(\mu)} + s_n;$$

zu seiner Abschätzung sind die in (3.5), (3.6) auftretenden Operationen zu analysieren.

Für die $g^{(\kappa)}(x_n)$ benötigt man gemäß (1.12) zunächst die Werte $k' \frac{\pi}{2N}$ für $k'=1, 2, \dots, N-1$. Hierbei treten jeweils 2 Gleitkomma-Operationen auf, ferner

liegt π nur gerundet vor. Es folgt

$$(3.8) \quad \mathbf{gl} \left(k' \frac{\pi}{2N} \right) = k' \frac{\pi}{2N} (1 + \delta), \quad (1 - \tau)^3 \leq 1 + \delta \leq (1 + \tau)^3$$

und daher auf Grund des Mittelwertsatzes

$$(3.9) \quad \sin \left(\mathbf{gl} \left(k' \frac{\pi}{2N} \right) \right) = \sin \left(k' \frac{\pi}{2N} \right) + D, \quad |D| \leq |\delta| \frac{\pi}{2} < 5\tau.$$

Die Koeffizienten in

$$\frac{h^\kappa}{\kappa!} g^{(\kappa)}(x_n) = (-1)^{\frac{\kappa-1}{2}} \sum_{k=1}^l (2t_k) \frac{(2kh)^\kappa}{\kappa!} \sin(2kx_n) \quad (\kappa \text{ ungerade})$$

berechnet man durch sukzessive Multiplikation von $(2t_k)$ mit den Faktoren

$$2k \frac{h}{j} = 2k \cdot \frac{\pi}{2Nj} \quad (j=1, 2, \dots, \kappa),$$

die zugehörigen numerischen Werte sind wie in (3.8) darstellbar. Es folgt

$$\mathbf{gl} \left(2t_k \frac{(2kh)^\kappa}{\kappa!} \right) = 2t_k \frac{(2kh)^\kappa}{\kappa!} (1 + \varepsilon_{k, \kappa}), \quad (1 - \tau)^{4\kappa} \leq 1 + \varepsilon_{k, \kappa} \leq (1 + \tau)^{4\kappa}.$$

Diese Ausdrücke sind mit $\mathbf{gl}(\sin(2kx_n))$ zu multiplizieren und zu summieren, wobei (3.2), (3.3) und (3.9) zu beachten ist. Wir erhalten so eine Darstellung

$$(3.10) \quad \mathbf{gl} \left(\frac{h^\kappa}{\kappa!} g^{(\kappa)}(x_n) \right) = (-1)^{\frac{\kappa-1}{2}} \sum_{k=1}^l (2t_k) \frac{(2kh)^\kappa}{\kappa!} (\sin(2kx_n) + D_k)(1 + \eta_{k, \kappa})$$

mit $|D_k| \leq 5\tau$ und – vgl. [5], Bd. 1, S. 19 –

$$|\eta_{k, \kappa}| \leq (4\kappa + l + 1) \tau'.$$

Wir bezeichnen

$$(3.11) \quad \mathbf{gl} \left(\frac{h^\kappa}{\kappa!} g^{(\kappa)}(x_n) \right) = : \frac{h^\kappa}{\kappa!} \tilde{g}_\kappa(x_n) = : \frac{h^\kappa}{\kappa!} (g^{(\kappa)}(x_n) + d_\kappa(x_n)).$$

Dann ist nach (3.10)

$$d_\kappa(x_n) = (-1)^{\frac{\kappa-1}{2}} \sum_{k=1}^l (2t_k) (2k)^\kappa [(\sin(2kx_n) + D_k) \eta_{k, \kappa} + D_k]$$

darstellbar und wegen $|\sin(2kx_n) + D_k| \leq 1$ – vgl. (3.9)! – durch

$$(3.12) \quad |d_\kappa(x_n)| \leq [(4\kappa + l + 1) \tau' + 5\tau] F_\kappa \leq (4\kappa + l + 6) \tau' F_\kappa$$

abschätzbar.

Wenn wir die Bezeichnungen (3.11) auch für geradzahliges κ verwenden, so ergibt sich völlig analog zu den bisherigen Überlegungen die Ungleichung (3.12) allgemeiner für alle $\kappa \geq 1$; im Fall $\kappa = 0$ erhält man

$$(3.12') \quad |d_0(x_n)| \leq (l+7) \tau' F_0.$$

Bei der Auswertung von (3.5) beachten wir zunächst, daß für $\mu = 1, 2, 3, \dots$

$$\mathbf{gl} \left(\frac{h}{\mu} \right) = \mathbf{gl} \left(\frac{\pi}{2N\mu} \right) = \frac{h}{\mu} (1 + \xi_\mu), \quad (1-\tau)^2 \leq 1 + \xi_\mu \leq (1+\tau)^2$$

gilt. Es folgen Gleichungen der Form

$$\begin{aligned} \tilde{u}_n^{(0)} &= \tilde{u}_n, & h\tilde{u}_n^{(1)} &= h\tilde{u}'_n(1 + \rho_1), \\ \frac{h^\mu}{\mu!} \tilde{u}_n^{(\mu)} &= \frac{h}{\mu} \frac{h}{\mu-1} \sum_{\kappa=0}^{\mu-2} \frac{h^{\mu-2-\kappa}}{(\mu-2-\kappa)!} \tilde{g}_{\mu-2-\kappa}(x_n) \frac{h^\kappa}{\kappa!} \tilde{u}_n^{(\kappa)}(1 + \rho_{\mu, \kappa}) \quad (\mu=2, 3, \dots) \end{aligned}$$

mit

$$|\rho_1| \leq 3\tau', \quad |\rho_{\mu, \kappa}| \leq (\mu+4) \tau' \quad (\mu=2, 3, \dots)$$

bzw. nach elementarer Umformung

$$\begin{aligned} \tilde{u}_n^{(0)} &= \tilde{u}_n, & \tilde{u}_n^{(1)} &= \tilde{u}'_n(1 + \rho_1), \\ \tilde{u}_n^{(\mu)} &= \sum_{\kappa=0}^{\mu-2} \binom{\mu-2}{\kappa} \tilde{g}_{\mu-2-\kappa}(x_n) \tilde{u}_n^{(\kappa)}(1 + \rho_{\mu, \kappa}) \quad (\mu=2, 3, \dots). \end{aligned}$$

Da die $\tilde{u}_n^{(\mu)}$ der Rekursion (1.10) zu den Anfangswerten $\tilde{u}_n, \tilde{u}'_n$ genügen, ergeben sich für die in (3.4) definierten $\varepsilon_n^{(\mu)}$ die Beziehungen

$$(3.13) \quad \begin{aligned} \varepsilon_n^{(0)} &= 0, & \varepsilon_n^{(1)} &= \tilde{u}'_n \rho_1, \\ \varepsilon_n^{(\mu)} &= \sum_{\kappa=0}^{\mu-2} \binom{\mu-2}{\kappa} [\tilde{g}^{(\mu-2-\kappa)}(x_n) \varepsilon_n^{(\kappa)} \\ &\quad + \{\tilde{g}_{\mu-2-\kappa}(x_n) \rho_{\mu, \kappa} + d_{\mu-2-\kappa}(x_n)\} \tilde{u}_n^{(\kappa)}] \quad (\mu=2, 3, \dots, p). \end{aligned}$$

Zum Zweck einer a posteriori Abschätzung der $\varepsilon_n^{(\mu)}$ verwenden wir die während der Lösung der Differentialgleichung zu bestimmenden Größen

$$\frac{h^\kappa}{\kappa!} U^{(\kappa)} := \max_{n=0}^{N-1} \left| \frac{h^\kappa}{\kappa!} \tilde{u}_n^{(\kappa)} \right| \quad (\kappa=0, 1, 2, \dots).$$

Ferner beachten wir, daß nach (3.12), (3.12') für $0 \leq \kappa \leq \mu-2, 2 \leq \mu \leq p+1$

$$|d_{\mu-2-\kappa}(x_n)| \leq (4\mu+l-1) \tau' F_{\mu-2-\kappa} \leq 0,1 \cdot F_{\mu-2-\kappa}$$

und daher

$$\begin{aligned} |\tilde{g}_{\mu-2-\kappa}(x_n) \rho_{\mu, \kappa} + d_{\mu-2-\kappa}(x_n)| &\leq [1,1(\mu+4) + (4\mu+l-1)] \tau' F_{\mu-2-\kappa} \\ &\leq (5,1\mu+l+3,4) \tau' F_{\mu-2-\kappa} \end{aligned}$$

gilt. Demgemäß definieren wir Größen $E_\mu \in \mathbb{R}$ durch die Rekursion

$$(3.14) \quad \begin{aligned} E_0 &= 0, & E_1 &= 3 \tau' U^{(1)}, \\ E_\mu &= \sum_{\kappa=0}^{\mu-2} \binom{\mu-2}{\kappa} F_{\mu-2-\kappa} [E_\kappa + (5,1\mu+l+3,4) \tau' U^{(\kappa)}] \quad (\mu=2, 3, \dots, p+1). \end{aligned}$$

Dann lassen sich aus (3.13) induktiv die Ungleichungen

$$(3.15) \quad |\varepsilon_n^{(\mu)}| \leq E_\mu \quad (\mu=0, 1, \dots, p, n=0, 1, \dots, N-1)$$

erschließen.

Die ebenfalls in (3.4) definierten $\hat{\varepsilon}_n^{(\mu)}$ genügen Gleichungen der Form

$$(3.13') \quad \begin{aligned} \hat{\varepsilon}_n^{(1)} &= 0, \\ \hat{\varepsilon}_n^{(\mu)} &= \sum_{\kappa=0}^{\mu-2} \binom{\mu-2}{\kappa} [g^{(\mu-2-\kappa)}(x_n) \varepsilon_n^{(\kappa)} \\ &\quad + \{\tilde{g}_{\mu-2-\kappa}(x_n) \hat{\rho}_{\mu,\kappa} + d_{\mu-2-\kappa}(x_n)\} \tilde{u}_n^{(\kappa)}] \quad (\mu=2, 3, \dots, p+1), \end{aligned}$$

wobei die

$$|\hat{\rho}_{\mu,\kappa}| \leq (\mu+2) \tau'$$

abschätzbar sind. Demnach setzen wir

$$(3.14') \quad \begin{aligned} \hat{E}_1 &= 0, \\ \hat{E}_\mu &= \sum_{\kappa=0}^{\mu-2} \binom{\mu-2}{\kappa} F_{\mu-2-\kappa} [E_\kappa + (5,1\mu+l+1,4) \tau' U^{(\kappa)}] \quad (\mu=2, 3, \dots, p+1) \end{aligned}$$

und folgern aus (3.13'), (3.15) unmittelbar

$$(3.15') \quad |\hat{\varepsilon}_n^{(\mu)}| \leq \hat{E}_\mu \quad (\mu=1, 2, \dots, p+1, n=0, 1, \dots, N-1).$$

Die Summen in (3.6) werden zweckmäßigerweise von rechts her, d.h. mit $\mu=p$ beginnend, ausgewertet. Dann gilt, wie man mittels (3.2) in bekannter Weise erschließt, eine Darstellung

$$\tilde{u}_{n+1} = \sum_{\mu=0}^p \frac{h^\mu}{\mu!} \tilde{u}_n^{(\mu)} (1 + \eta_\mu)$$

mit

$$|\eta_0| \leq \tau < \tau', \quad |\eta_\mu| \leq \mu \cdot \tau' \quad (\mu=1, 2, \dots, p)$$

und folglich

$$\tilde{u}_{n+1} = \sum_{\mu=0}^p \frac{h^\mu}{\mu!} \tilde{u}_n^{(\mu)} + \left(\sum_{\mu=0}^p \frac{h^\mu}{\mu!} \tilde{u}_n^{(\mu)} \eta_\mu + \sum_{\mu=0}^p \frac{h^\mu}{\mu!} \varepsilon_n^{(\mu)} \right).$$

Der in der Klammer stehende Ausdruck ist gemäß (3.7) gleich σ_n ; somit erhalten wir die Abschätzung

$$(3.16) \quad |\sigma_n| \leq \sigma := \tau' U^{(0)} + \sum_{\mu=1}^p \frac{h^\mu}{\mu!} (\mu U^{(\mu)} \tau' + E_\mu) \quad (n=0, 1, \dots, N-1).$$

Wenn man die zweite Komponente in (3.6) analog behandelt, so ergeben sich die Ungleichungen

$$(3.16') \quad |\sigma'_n| \leq \sigma' := \tau' \hat{U}^{(1)} + \sum_{\mu=1}^p \frac{h^\mu}{\mu!} (\mu \hat{U}^{(\mu+1)} \tau' + \hat{E}_{\mu+1}) \quad (n=0, 1, \dots, N-1),$$

wobei für $\kappa=1, 2, \dots, p+1$ die

$$\hat{U}^{(\kappa)} := \max_{n=0}^{N-1} |\hat{u}_n^{(\kappa)}| \quad (\approx U^{(\kappa)})$$

definiert sind.

4. Abschätzung des globalen Fehlers

Es sei y die Lösung der Differentialgleichung (1.3), und zwar mit der Anfangsbedingung $z_0 = e_1$ oder e_2 . Wir bezeichnen für $n=0, 1, \dots, N$

$$\begin{aligned} y_n &:= y(x_n), & y_n^{(\mu)} &:= y^{(\mu)}(x_n) \quad (\mu=0, 1, 2, \dots), \\ z_n &:= z(x_n), & z_n^{(\mu)} &:= z^{(\mu)}(x_n) \quad (\mu=0, 1, 2, \dots), \end{aligned}$$

ferner seien die \tilde{v}_n wie in (3.6) definiert. Gesucht sind komponentenweise Abschätzungen der Vektoren

$$(4.1) \quad f_n := \begin{pmatrix} |\tilde{u}_n - y_n| \\ |\tilde{u}'_n - y'_n| \end{pmatrix} \quad (n=0, 1, \dots, N).$$

Nach (1.6), (2.18) bzw. (3.7) gelten für $n=0, 1, \dots, N-1$ die Beziehungen

$$z_{n+1} = z_n + \sum_{\mu=1}^p \frac{h^\mu}{\mu!} z_n^{(\mu)} + r_n, \quad \tilde{v}_{n+1} = \tilde{v}_n + \sum_{\mu=1}^p \frac{h^\mu}{\mu!} \tilde{v}_n^{(\mu)} + s_n.$$

Da die Komponenten $\tilde{v}_n^{(\mu)}$ der $\tilde{v}_n^{(\mu)}$ der Rekursion (1.10) mit den Anfangswerten $\tilde{u}_n, \tilde{u}'_n$ genügen, folgt aus Hilfssatz (2.10) für $\mu=0, 1, 2, \dots$

$$\tilde{v}_n^{(\mu)} = \begin{pmatrix} \alpha_\mu^1(x_n) & \alpha_\mu^2(x_n) \\ \alpha_{\mu+1}^1(x_n) & \alpha_{\mu+1}^2(x_n) \end{pmatrix} \tilde{v}_n =: A_n^{(\mu)} \tilde{v}_n$$

und daher

$$(4.2) \quad \tilde{v}_{n+1} = \left(I + \sum_{\mu=1}^p \frac{h^\mu}{\mu!} A_n^{(\mu)} \right) \tilde{v}_n + s_n.$$

Ebenfalls aus Hilfssatz (2.10) erschließt man

$$(4.3) \quad z_{n+1} = \left(I + \sum_{\mu=1}^p \frac{h^\mu}{\mu!} A_n^{(\mu)} \right) z_n + r_n.$$

Somit ergibt sich

$$(4.4) \quad \begin{aligned} v_0 - z_0 &= 0, \\ \tilde{v}_{n+1} - z_{n+1} &= \left(I + \sum_{\mu=1}^p \frac{h^\mu}{\mu!} A_n^{(\mu)} \right) (\tilde{v}_n - z_n) + (s_n - r_n) \\ &\quad (n=0, 1, \dots, N-1). \end{aligned}$$

Wegen (2.16) lassen sich die auftretenden Matrizen durch

$$Q := I + \sum_{\mu=1}^p \frac{h^\mu}{\mu!} \begin{pmatrix} a_\mu^1 & a_\mu^2 \\ a_{\mu+1}^1 & a_{\mu+1}^2 \end{pmatrix}$$

komponentenweise abschätzen. Wir definieren

$$s := \begin{pmatrix} \sigma \\ \sigma' \end{pmatrix}$$

wie in (3.16), (3.16') und übernehmen r aus (2.19). Dann folgen für die f_n die Ungleichungen

$$(4.5) \quad \begin{aligned} f_0 &= 0, \\ f_{n+1} &\leq Q f_n + (s+r) \quad (n=0, 1, \dots, N-1) \end{aligned}$$

und weiter, wenn wir

$$(4.6) \quad G := \sum_{\kappa=0}^{N-1} Q^\kappa$$

definieren,

$$(4.7) \quad f_n \leq \left(\sum_{\kappa=0}^{n-1} Q^\kappa \right) (s+r) \leq G(s+r) \quad (n=1, 2, \dots, N).$$

Da (im Fall $g \neq 0$) alle Koeffizienten von Q positiv sind, besitzt Q zwei verschiedene Eigenwerte $\mu_1, \mu_2 \in \mathbb{R}$; folglich existiert eine reelle, invertierbare Matrix C mit der Eigenschaft

$$C^{-1}QC = D := \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}.$$

D und C lassen sich elementar berechnen; hiermit wird

$$(4.8) \quad G = C \left(\sum_{\kappa=0}^{N-1} D^{\kappa} \right) C^{-1}.$$

5. Ergänzungen und Beispiele

Zur Bestimmung der Schrittweite h und der Ordnung p des Verfahrens ist folgendes Vorgehen zu empfehlen: man wählt ein geeignet erscheinendes h und ermittelt hierzu eine passende Ordnung p des Verfahrens durch eine a priori Abschätzung des lokalen Verfahrensfehlers, so wie in Abschnitt 2 beschrieben. Die dort erwähnte Konstante ε sollte nicht wesentlich kleiner als die Rechengenauigkeit τ sein, da andernfalls die letzten Summanden in (1.9) ebenfalls betragsmäßig kleiner als τ sind und daher das numerische Ergebnis kaum beeinflussen. Die besten Ergebnisse und die günstigsten Rechenzeiten werden erfahrungsgemäß dann erreicht, wenn das so ermittelte p zwischen 10 und 40 liegt; andernfalls sollte man h vergrößern bzw. verkleinern. Bisher hat sich die Wahl

$$(5.1) \quad h = \frac{\pi}{2N}, \quad N \approx 5 \cdot \max \{1, \sqrt{|\lambda|}\}$$

gut bewährt.

Die folgenden Zahlenbeispiele wurden im Rechenzentrum der Universität Regensburg auf einer TR 440 mit der eingebauten doppeltgenauen Arithmetik (die etwa einer 25stelligen Dezimal-Arithmetik entspricht) gerechnet.

Beispiel I (Hill [1])	Beispiel II
$l = 3$	$l = 10$
$\lambda = 1,1588439396$	$\lambda = 17,2$
$t_1 = -0,05704401875$	$t_k = \frac{1}{k^2} \quad (k = 1, 2, \dots, 10)$
$t_2 = 0,00038323800$	
$t_3 = -0,00000917329$	
$\varepsilon = 10^{-19}$	$\varepsilon = 10^{-19}$
$N = 6$	$N = 24$

Die errechneten Werte sind in der folgenden Tabelle ausgedruckt; aus Platzgründen sind die Lösungen der Differentialgleichungen hier nur mit ca. 10 Dezimalstellen, die Fehlerschranken mit 3 Dezimalstellen angegeben. Von den Abschätzungen der lokalen Rundungsfehler s , die für y_1, y_2 getrennt berechnet wurden, sind hier nur die jeweils größeren Werte berücksichtigt. Bei Abschätzung des Gesamtfehlers gemäß (4.7) wurde

$$\tau' \leq 0,5 \cdot 10^{-24}$$

benutzt.

	Beispiel I	Beispiel II
p	21	23
$y_1 \left(\frac{\pi}{2} \right)$	-0,0771302845	1,0434199068
$y_1' \left(\frac{\pi}{2} \right)$	-1,0706105528	-0,9771279472
$y_2 \left(\frac{\pi}{2} \right)$	0,9222866582	0,0509126183
$y_2' \left(\frac{\pi}{2} \right)$	-0,1632325971	0,9107089596
r	$\begin{pmatrix} 0,213 \cdot 10^{-20} \\ 1,89 \cdot 10^{-20} \end{pmatrix}$	$\begin{pmatrix} 0,126 \cdot 10^{-20} \\ 5,57 \cdot 10^{-20} \end{pmatrix}$
$s \leq$	$\begin{pmatrix} 3,04 \\ 7,79 \end{pmatrix} \cdot \tau'$	$\begin{pmatrix} 3,60 \\ 44,9 \end{pmatrix} \cdot \tau'$
G	$\begin{pmatrix} 8,81 & 4,89 \\ 6,41 & 8,83 \end{pmatrix}$	$\begin{pmatrix} 1,84 \cdot 10^3 & 0,404 \cdot 10^3 \\ 8,39 \cdot 10^3 & 1,85 \cdot 10^3 \end{pmatrix}$
$f_N \leq$	$\begin{pmatrix} 0,111 \cdot 10^{-18} \\ 0,180 \cdot 10^{-18} \end{pmatrix}$	$\begin{pmatrix} 0,248 \cdot 10^{-16} \\ 1,13 \cdot 10^{-16} \end{pmatrix}$
$ \tilde{v} - v \leq$	$0,912 \cdot 10^{-19}$	$0,452 \cdot 10^{-16}$

Die in der letzten Zeile der Tabelle angegebenen Fehlerschranken für die numerisch gewonnenen Näherungen \tilde{v} von v resultieren aus (1.2) – bzw. in Beispiel II aus (1.2') – mittels einer elementaren Fehleranalyse, in der die angegebenen Abschätzungen für f_N eingehen. Die berechneten Werte \tilde{v} , auf 19 Dezimalstellen gerundet, sind

$$\tilde{v} = 0,9284167225828297331$$

bzw. in Beispiel II

$$\tilde{v} = 0,1431980134051061052.$$

Wie Vergleichsrechnungen mit erhöhter Genauigkeit zeigen, liegen die tatsächlichen Fehler der \tilde{v} bei $0,1 \cdot 10^{-19}$ bzw. $0,5 \cdot 10^{-20}$; somit sind die berechneten Fehlerschranken um etwa den Faktor 10 bzw. 10^4 zu grob. Bei nicht zu großen Parametern liefert die hier entwickelte Fehlerrechnung also durchaus realistische Schranken.

Entscheidend hierfür ist die Behandlung der Fehlerfortpflanzung gemäß Abschnitt 4. Mittels der sonst üblichen Fehleranalyse – vgl. z.B. [9], S. 106 ff. – erhält man an Stelle von (4.7) bezüglich der Maximumsnorm im \mathbb{R}^2 die Abschätzungen

$$(5.2) \quad \|f_n\| \leq \gamma(\|s\| + \|r\|) \quad (n = 1, 2, \dots, N),$$

wobei

$$(5.3) \quad \gamma = \frac{1}{hM} \left(\exp \left(M \frac{\pi}{2} \right) - 1 \right),$$

$$M = \frac{1}{h} \cdot \max \left\{ \sum_{\mu=1}^p \frac{h^\mu}{\mu!} (a_\mu^1 + a_\mu^2), \sum_{\mu=1}^p \frac{h^\mu}{\mu!} (a_{\mu+1}^1 + a_{\mu+1}^2) \right\}.$$

Hier ist

$$\gamma = 24,2 \quad \text{in Beispiel I,}$$

$$\gamma = 0,484 \cdot 10^{15} \quad \text{in Beispiel II;}$$

insbesondere im letzteren Fall ist (5.2) im Gegensatz zu (4.7) kaum brauchbar.

Das hier beschriebene Taylor-Verfahren wurde mit zahlreichen anderen Verfahren der numerischen Integration verglichen; es war hinsichtlich der erreichbaren Genauigkeit und des geringen Rechenaufwandes allen anderen Verfahren überlegen. Als die zweitbeste Methode erwies sich das Extrapolationsverfahren nach Gragg, Bulirsch und Stoer, wie es beispielsweise in [9], S. 145 ff. beschrieben ist. Hierbei ist es vom Lösungsverhalten der Differentialgleichung her sinnvoll, eine konstante Grundschriftweite

$$H = \frac{\pi}{2N}$$

vorzugeben und dann zu jedem Teilintervall die Anzahl p der Schrittweiten h_i , d.h. die Spaltenzahl des Extrapolationsschemas, mittels eines geeigneten Genauigkeitskriteriums vom Programm aus zu steuern. In den angegebenen Beispielen wurde $N = 10$ bzw. $N = 20$ gewählt; die zugehörigen Werte von p lagen bei 6 bzw. bei 8. Mit dem Extrapolationsverfahren wurden die charakteristischen Exponenten ebenso genau wie mit dem Taylor-Verfahren berechnet, jedoch ist eine entsprechende zuverlässige Fehlerrechnung bisher nicht bekannt. Die folgende Tabelle enthält die Rechenzeiten, die zur Bestimmung von ν mit einer (tatsächlichen) Genauigkeit von 19 Dezimalstellen benötigt wurden; zusätzlich wurde beim Taylor-Verfahren die für die Fehleranalyse benötigte Rechenzeit gemessen.

	Beispiel I	Beispiel II
Taylor-Verfahren	0,339 s	1,405 s
Fehlerrechnung	0,075 s	0,144 s
Extrapolation	2,406 s	21,49 s

Im Fall der Mathieschen Differentialgleichung ($l=1$) wird der charakteristische Exponent noch schneller mittels der von Wagenführer [10] angegebenen Determinantenmethode berechnet; die Rechenzeiten betragen durchweg weniger

als die Hälfte der beim Taylorverfahren benötigten Zeiten. Die auch im Fall $l > 1$ anwendbare Determinantenmethode nach Mennicken [4] soll in Kürze numerisch diskutiert und mit dem Taylor-Verfahren verglichen werden.

Literatur

1. Hill, G.W.: On the part of the motion of the lunar perigee, which is a function of the mean motions of the sun and the moon. *Acta math.* **8**, 1–36 (1886)
2. Ince, E.L.: *Ordinary Differential Equations*. Nachdruck. New York: Dover Publications 1956
3. Meixner, J., Schäfke, F.W.: *Mathiesche Funktionen und Sphäroid-Funktionen*. Berlin-Heidelberg-New York: Springer 1954
4. Mennicken, R.: On the Convergence of Infinite Hill-Type Determinants. *Arch. for Rat. Mech. and Anal.* **30**, 12–37 (1968)
5. Mennicken, R., Wagenführer, E.: *Numerische Mathematik 1, 2*. Reinbek: rororo-Vieweg 1976, 1977
6. Schäfke, F.W.: Ein Verfahren zur Berechnung des charakteristischen Exponenten der Mathieschen Differentialgleichung. I. *Numer. Math.* **3**, 30–36 (1960)
7. Schäfke, F.W., Ebert, R., Groh, H.: Ein Verfahren zur Berechnung des charakteristischen Exponenten der Mathieschen Differentialgleichung. II. *Numer. Math.* **4**, 1–7 (1962)
8. Schäfke, F.W., Schmidt, D.: Ein Verfahren zur Berechnung des charakteristischen Exponenten der Mathieschen Differentialgleichung. III. *Numer. Math.* **8**, 68–71 (1966)
9. Stoer, J., Bulirsch, R.: *Einführung in die Numerische Mathematik II*. Berlin, Heidelberg, New York: Springer 1973
10. Wagenführer, E.: Ein Verfahren höherer Konvergenzordnung zur Berechnung des charakteristischen Exponenten der Mathieschen Differentialgleichung. *Numer. Math.* **27**, 53–65 (1976)

Eingegangen am 21. März 1978